PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 4/00, C12N 1/00, 5/00, 15/09, 15/12, C12Q 1/68, G01N 33/566

A1

(11) International Publication Number: **WO 95/24419**

(43) International Publication Date: 14 September 1995 (14.09.95)

(21) International Application Number: PCT/US95/03208

(22) International Filing Date: 13 March 1995 (13.03.95)

(30) Priority Data:

08/209,835	11 March 1994 (11.03.94)	US
08/369,832	6 January 1995 (06.01.95)	US

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.

(71) Applicant *(for all designated States except US)*: ARIAD PHARMACEUTICALS, INC. [US/US]; 26 Landsdowne Street, Cambridge, MA (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: RICKLES, Richard, J. [US/US]; 35 Eustis Street, Cambridge, MA 02138 (US). BRUGGE, Joan, S. [US/US]; 50 Sarah Way, Concord, MA 01742 (US). BOTFIELD, Martyn, C. [US/US]; 44 Vernon Street, Brookline, MA 02146 (US). ZOLLER, Mark, J. [US/US]; 7 Sutton Place, Weston, MA 02193 (US).

(74) Agent: BERSTEIN, David, L.; ARIAD Pharmaceuticals, Inc., 26 Landsdowne Street, Cambridge, MA 02139 (US).

(54) Title: METHODS AND MATERIALS FOR IDENTIFYING INHIBITORS OF MOLECULAR INTERACTIONS MEDIATED BY SH3 DOMAINS

(57) Abstract

This disclosure concerns materials and methods for identifying inhibitors of molecular interactions mediated by SH3 domains, and for compositions and uses relating to such inhibitors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

Methods and Materials for Identifying Inhibitors Of Molecular Interactions Mediated by SH3 Domains

Technical Field

This invention relates to materials and methods for identifying inhibitors of molecular interactions mediated by SH3 domains, and for compositions and uses relating to such inhibitors.

Introduction

A fundamental goal of recent medical and pharmaceutical research has been the elucidation of mechanisms of diseases and the identification of critical points within such mechanisms for effective pharmaceutical intervention. Unfortunately, in some cases the disease mechanism has not yet been sufficiently defined to permit precise intervention, while in other cases gaps in technology have hampered the research effort.

One mechanistic aspect of cellular function in both normal and disease states which has attracted increasing attention is cellular signal transduction, the series of events leading from extracellular events to intracellular sequelae. Numerous proteins that function as signal transducing molecules have been identified. These include receptor and non-receptor tyrosine kinases, phosphatases and other molecules with enzymatic or regulatory activities. One feature common to many of these molecules is the capacity to associate specifically with other proteins to form a signaling complex that can alter cell activity.

Signaling proteins often contain domain(s) of conserved sequence, which serve as non-catalytic modules that direct protein-protein interactions during signal transduction. Two such domains have been termed the src homology domain 2 (SH2) and src homology domain 3 (SH3). These domains are relatively small (~100 amino acids for SH2 and ~ 60 amino acids for SH3) and are found in various combinations and locations in different proteins. For example, some members of the src-family of tyrosine kinases, *e.g.*, Abl, GRB2 and PI3K each contain both SH2 and SH3 domains. In contrast, the tyrosine kinase Syk has two SH2 domains, while two components of NADPH oxidase, p47-phox and p67-phox, each have two SH3 domains. The presence of multiple SH2 and/or SH3 domains within a protein increases the variety of potential protein-protein interactions.

Aspects of the structures of various SH2 domains are known and certain aspects of their role in signal transduction is becoming better understood. SH2 domains direct the association of specific proteins by binding selectively and with specificity to protein sequences containing phosphotyrosine. For example, upon ligand binding, the PDGF ß-receptor dimerizes and autophosphorylates multiple tyrosine residues. This phosphorylation triggers the physical association of SH2-containing proteins such as c-src, PLC-gamma, PI3K and ras-GAP with the receptor, forming a signaling complex. An analysis of SH2 binding using natural ligands containing mutations at residues surrounding the site of phosphorylation as well as a screen of combinatorial peptide libraries using SH2 domains has revealed the specificity of SH2 binding.

Less is known about ligand specificity for SH3 domains. Initial inroads towards understanding SH3 ligand specificity came first from a screen of a cDNA expression library using the c-abl SH3 domain as a probe. Two proteins were identified, termed 3BP-1 and 3BP-2. Short proline-rich sequences within these proteins were demonstrated as important for c-abl SH3 binding.

Pharmaceutical agents which interfere with the formation or stability of signaling complexes formed by proteins containing one or more SH3 domains and their natural ligands could be used to treat or prevent the diseases or their pathological effects mediated by such complexes—but only if these interfering agents have sufficient binding specificity. Unfortunately, too little has been known about the natural ligands and the binding specificities of the various SH3 domains to know whether selective interference with SH3 binding would actually be possible, and if it were, to permit the identification or design of agents capable of so acting. Indeed, the limited apparent sequence requirements (proline rich sequences) for binding to SH3 domains together with the limited expected affinities of interactions therewith were considered discouraging, especially in comparison to emerging information for SH2 domains.

Summary of the Invention

The present inventors have discovered that protein-protein interactions mediated by SH3 domains are characterized by a surprisingly high level of binding specificity, far beyond a simple requirement for "proline rich" ligands. Indeed, the present inventors have found that peptide ligands (including "synthetic" peptide ligands, i.e., ligands containing amino acid sequences not heretofore found in nature) can be identified which bind to a particular SH3 domain selectively and with high affinity. They have developed a general methodology for identifying specific peptide ligands for SH3 domains and have applied it to representative SH3 domains for which they identified selective, high affinity ligands.

Using this technology for identifying specific ligands for SH3 domains, the present inventors have devised *in vitro* and *in vivo* assays for identifying inhibitors of SH3-mediated interactions. The materials and method used in those assays provide for high-throughput identification of SH3 inhibitors.

These discoveries thus permit, for the first time, the practical identification of compounds which specifically block the signaling of an SH3-bearing protein of interest and which may be used as pharmaceutical agents.

Interestingly, the present inventors' initial success in identifying specific binding partners for individual SH3 domains hinged on a serendipitous polymerase accident, discussed in greater detail below. That accident led to the first positive results demonstrating the feasibility of their overall approach. Those and subsequent results, combined with a novel extension of BIAcore® technology—which itself should be of widespread interest and applicability—confirmed the significance of their findings, and again, the feasibility of their overall approach.

This invention provides materials and methods, both *in vitro* and *in vivo*, for identifying substances which bind selectively to an SH3 domain of interest (i.e., to a protein or peptide, preferably of human origin or sequence, containing an SH3 domain of interest) or which selectively block or otherwise inhibit the interaction of proteins mediated by one or more SH3 domains.

This invention thus provides a reproducible, generally applicable method for obtaining an SSL which binds selectively to an SH3 domain of interest and further provides illustrative peptide SSLs for representative SH3 domains. The SSLs, fusion proteins containing an amino acid sequence corresponding to that of an SSL and DNA or RNA molecules encoding them are useful in the present methods for identifying SH3 binding agents and SH3 blocking agents.

SSLs may be identified by combining an SH3 domain of interest (*e.g.* in the form of a peptide, protein, protein fragment or fusion protein) with one or more candidate ligands [*e.g.* a peptide (including peptide SSLs which contain an amino acid sequence not present in a naturally occurring protein), peptoid, protein, protein fragment, fusion protein, non-peptide backbone displaying "side chains" of natural amino or imino acids, nucleic acid or any hybrid of the preceding] under conditions permitting a ligand to bind to an SH3 domain to form an SH3-ligand complex, then detecting which if any candidates 1) bound to the SH3 domain or 2) blocked binding the SH3 domain to a known SH3 ligand. This methodology may be applied to SH3 domains present in Abl, Src, Grb-2, Nck or p85 PI 3-K, as well as to SH3 domains from other proteins. A candidate ligand which bound to the SH3 domain may be recovered from the SH3-ligand complex and further characterized. This can be accomplished, *e.g.* by separating unbound candidate ligands from any SH3 domain-ligand complexes so

formed, dissociating the complexed ligand(s) from the SH3 domain-ligand complexes, enriching for the selected ligands by re-contacting the dissociated ligands with the SH3 domain under conditions permitting a ligand to bind to an SH3 domain to form a complex, and detecting which any candidate ligands bound to the SH3 domain. The enriched ligand(s) may be dissociated from the complexes so formed and may then be structurally characterized, if desired. Peptide ligands comprising an amino acid sequence not present in a naturally occurring protein are of special interest, especially in cases in which the SH3 domain is other than an *Abl* or p85 PI 3-Kinase SH3 domain. A ligand so selected may be re-assayed as above, but against different SH3 domains, to determine the presence and/or extent of binding of the ligand to a different SH3 domain—*i.e.*, to confirm the selectivity of the interaction with the SH3 domain against which it was identified. Peptide SSLs may thus be obtained containing 7 amino acid residues or more, and as disclosed below, may in some applications contain at least 10 amino acids residues, and in some cases as least 12, 14 or 15 amino acid residues. Peptide SSLs may be so obtained for SH3 domain-containing proteins include (1) members of the src-family protein tyrosine kinases (*Src, Lyn, Fyn, Lck, Hck, Fgr, Yes*), (2) *Grb*-2, which has two SH3 domains, (3) *Sprk*, a threonine/serine protein kinase, (4) *Tsk*, (5) *Btk*, (6) *Txk* and other *Tec* family members, (7) *Vav*, (8) GTPase Activating Protein (GAP), (9) p40, p47, and p67 proteins of the neutrophil oxidase complex, and (10) p85 phosphatidylinositol 3' kinase, (11) *Crk*, (12) phospholipase C gamma, and (13) *Abl*. Ligands containing peptide sequences not found in naturally occurring proteins may be selected for SH3 domains of *Src, Lyn, Fyn, Yes, Txk, Tsk, Btk*, NADPH oxidase p47-phox and p67-phox, *Crk* and other proteins. DNAs encoding such peptide ligands may be readily prepared using conventional methods and materials.

For example, a mixture of candidate ligands, preferably peptides, may be contacted with an SH3 domain of interest and incubated under conditions permitting the formation of an SH3-ligand complex. Uncomplexed candidate ligands and SH3-ligand complex(es) are separated from each other, and ligand(s) bound to the SH3 domain are separately recovered and identified. A preferred approach is disclosed *infra* in which the SH3 domain of interest, in the form of a fusion protein, is used to select peptide ligands from combinatorial libraries displayed on phage. Sequences so identified may be re-assayed as mentioned above, whether in the form of phage-displayed peptides, as part of any recombinant protein or as synthetic peptides.

SSLs so obtained may then be used as follows to identify SH3 binding or blocking agents. Generically stated, the method of this invention employs: (1) a peptide which contains an SH3 domain of interest, (2) a specific SH3 ligand (SSL) of this invention (which in *in vitro* embodiments can be a peptide or nonpeptide, but is

preferably not a protein) which is capable of selectively binding to the SH3 domain of interest to form an SH3-SSL complex and (3) a compound (the "test substance") to be evaluated for its ability to bind competitively to a selected SH3 domain, or block molecular interactions mediated by that SH3 domain, in a specific and selective manner. The terms "peptide" and "protein" are used interchangeably herein, except in those instances in which an explicit distinction has been drawn. The SSL may be a peptide SSL comprising a peptide sequence which is present in a naturally occurring protein or a peptide sequence which is *not* present in a naturally occurring protein. The method is carried out by combining the three components mentioned above, or compositions comprising them; incubating the resulting test mixture under conditions permitting the formation of an SH3-SSL complex; and measuring the ability of the test substance to compete with the SSL for binding to the SH3 domain or to otherwise block the formation or reduce the observed level of SH3-SSL complex. This method is a powerful and general method, and should be applicable to any SH3 domain and susceptible to variety of configurations, including both *in vitro* and *in vivo* formats. Depending on the specific assay configuration, it may be important to use known concentrations of SH3 domain, SSL and/or test substance. For comparative purposes, the assay may also be carried out in the absence of the test substance or in the presence of varying concentrations of test substance. One may carry out the measuring step by assaying for SH3-SSL complex, non-complexed SSL and/or non-complexed test substance or by measuring the occurence of an event mediated by the presence or formation of the SH3-SSL complex or an SH3-test substance complex.

The test substance may be present in a solution, referred to as a test solution. Alternatively, especially (but not exclusively) for *in vitro* assays, the test substance may be present in a test mixture comprising an emulsion, suspension or other mixture; exposed on the surface of a cell, virus, phage, *etc.*; or immobilized on a solid support.

In an *in vitro* format, a binding assay is conducted to identify a compound capable of binding to the SH3 domain in the presence of a specific SH3 ligand for that SH3 domain or otherwise capable of blocking the formation or reducing the observed level of SH3-SSL complex. In one embodiment, the binding assay is a competitive binding assay in which the three components are combined and incubated under conditions permitting the formation of an SH3-SSL complex. The ability of the test substance to bind to the selected SH3 domain or otherwise block the SH3-mediated interaction in the presence of the specific SH3 ligand is determined. Binding to the SH3 domain or otherwise blocking the SH3-mediated interaction may be measured directly or indirectly (*e.g.*, BIAcore® and other SPR technologies (*BIAtechnology Handbook*, Pharmacia Biosensor AB, Uppsala, Sweden, 1994), fluorescence anisotropy and allied technologies (Luminescent Spectroscopy of

Proteins, 164pp, E. A. Permyakov, CRC Press, Inc, Boca Raoton, FL, 1992), flow cytometry and allied technologies (*Flow Cytometry and Cell Sorting*, 223pp., A. Radbruch, ed., Springer-Verlag, New York, NY, 1992), ELISA, RIA and allied methodologies (*An Introduction to Radioimmunoassays and Related Techniques*, 290 pp., T. Chard, Elsevier Science Publishers, Amsterdam, The Netherlands, 1990), competitive and non-competive affininty interactions (*Immobilized Affinity Ligand Techniques*, 454 pp., G. T. Hermanson, A. K. Mallia and P.K. Smith, eds., Academic Press, Inc., San Diego, CA, 1992). If binding of the SH3 domain and the specific SH3 ligand occurs to a lesser extent in the presence of the test substance than in its absence, for instance, if the presence of the test substance reduces the concentration of SH3-SSL complex or increases the concentration of non-complexed SSL, then the test substance is an SH3 binding or blocking agent. If the structure of the SH3 binding agent so identified is not yet known, the compound may then be isolated from the other assay components and characterized. It may be re-evaluated, if desired, using one or more different SSLs to the same SH3 domain and/or assayed in similar binding assays with different SH3 domains to confirm the selectivity of the interaction with the SH3 domain with which it was identified. If desired, the binding of the SH3 binding agent to the SH3 domain with which it was identified may be characterized biochemically, *e.g.* through the use of BIAcore® technology, described in greater detail below. The SH3 binding agent so identified may be assayed in an *in vivo* SH3 blocking assay as described below and may further be evaluated for pharmacological activity in various *in vitro* and/or *in vivo* assays, as desired.

in vivo assays can be conducted in analogous manner using cells containing the SH3 domain of interest and a ligand therefor. The cells are cultured or maintained in a medium suitable for cell growth. The test substance is added to the cells, *e.g.* to the medium in which the cells are cultured, and the culture is incubated under conditions permitting formation of a complex between the SH3 domain of interest and its ligand. If binding of the SH3 domain and its ligand occurs to a lesser extent in the presence of the test substance than in its absence, for instance, if the presence of the test substance reduces the concentration of SH3-ligand complex or increases the concentration of non-complexed ligand, then the test substance is an SH3 binding or blocking agent. The presence or absence of SH3-ligand complex may be measured directly or indirectly (*e.g.*, by measuring the occurence of an event mediated by the presence or formation of the SH3-ligand complex or an SH3-test substance complex).

An illustrative *in vivo* format relies upon genetically engineered cells capable of expressing a reporter gene under SH3-mediated transcriptional control. These cells contain an SH3 domain of interest and the corresponding SSL, each in the form of a distinct fusion protein. Each such fusion protein thus comprises, among other

component regions, at least one SH3 domain or SSL sequence. The SH3 fusion protein and SSL fusion protein are capable of forming a complex with each other. The cells express the reporter gene unless a substance is present which binds to the relevant SH3 domain or otherwise blocks the SH3-mediated interaction required for transcription of the reporter gene. In this assay, the cells are cultured or maintained in a suitable culture medium to establish a base-line for expression of the reporter gene. The test substance is added to the culture medium and the ability of the test substance to inhibit expression of the reporter gene is measured. If the level of reporter gene expression is reduced in the presence of the test substance, the test substance is an SH3 blocker with respect to the SH3 domain involved in transcriptional control. If the structure of the SH3 blocking agent so identified is not yet known, the compound may then be isolated from the other assay components and characterized. It may be re-evaluated, if desired, using engineered cells containing a fusion protein with one or more different SSLs to the same SH3 domain and/or assayed with cells containing fusion proteins comprising different SH3 domains and corresponding SSL sequences to confirm the selectivity of the interaction with the SH3 domain with which it was identified. If desired, the binding affiinity of the SH3 blocking agent for the SH3 domain with which it was identified may be determined, *e.g.* such as through the use of BIAcore® technology, described in greater detail below. The SH3 blocking agent so identified may be assayed in an *in vitro* SH3 binding assay as described above and may further be evaluated for pharmacological activity in various *in vitro* and/or *in vivo* assays, as desired.

Specific SH3 ligands and SH3 binding or blocking agents identified by the present method can be identified from peptide libraries as well as from test substances obtained from a wide variety of sources including, *e.g.*, microbial broths; cellular extracts; conditioned media from cell lines or from host cells transformed with genetic libraries; collections of synthetic compounds; combinatorial libraries or synthetic programs based on conventional medicinal chemistry approaches or structure-based drug design.

This invention thus provides a means for identifying selective SH3 binding or blocking agents. As noted at the outset, SH3 domains are present in a wide variety of proteins, including proteins involved in intracellular signal transduction pathways, proteins associated with the cytoskeleton and adapter proteins. Accordingly, SSLs and specific SH3 binding or blocking agents identified through this invention may be useful for a variety of purposes. First, they may be useful as biological reagents in assays as described herein for functional classification of an SH3 domain of a particular protein, particularly a newly discovered protein. Families or classes of

SH3-bearing proteins may now be defined functionally, with respect to ligand specificity.

Moreover, SH3 binding or blocking agents of this invention can be used to inhibit the occurrence of biological events resulting from molecular interactions mediated by an SH3 domain. This invention thus provides a method and reagents for inhibiting (totally or partially) the interaction between a protein containing an SH3 domain and a natural ligand thereto (i.e., a protein which normally binds in a cell to the SH3-bearing protein) or a biological activity mediated by such interaction. In this method, an SH3 binding or blocking agent, such as an SSL or an SH3 binder or blocker identified by this invention, including a peptide of 4-50 amino acid residues, more preferably 7-15 amino acid residues in length, is combined or contacted with the SH3 domain-containing protein, such as by introducing the SH3 binding or blocking agent into a cell in which the SH3-mediated interaction is to be inhibited. Following introduction of the SH3 binding or blocking agent, the interaction of the SH3 domain-bearing protein and its natural ligand is inhibited as may be readily detected. Inhibiting such interactions can be useful in research aimed at better understanding the biology of SH3-mediated events.

Such SH3 binding or blocking agents would be useful, for example, in the diagnosis, prevention or treatment of conditions or diseases resulting from a cellular processes mediated by an SH3-based interaction. For example, a patient can be treated to prevent the occurence or progression of osteoporosis, osteomalacia, or Paget's disease, or to reverse the course of these diseases by administering to the patient in need thereof an SH3 binding or blocking agent which selectively binds src SH3. There are many other conditions for which SH3 binding or blocking agents can be used therapeutically, including restenosis, rheumatoid arthritis, gout, asthma, emphysema, immune vasculitis, ulcerative colitis, psoriasis and acute respiratory distress syndrome, in which an SH3 of neutrophil oxidase p47 and p67 complex has been implicated. Other relevant conditions include chronic myelogenous leukemia, in which case SH3 domains of Grb-2 are targeted. It has recently been shown that the BCR-abl oncogene in CML participates in the ras pathway for growth stimulation through its interaction with Grb-2. In these cells, inhibition of the interaction of Grb-2 SH3 domains with the SOS oncogene will block its ability to stimulate cell proliferation. Still other relevant conditions include cancers such as breast cancer, glioblastomas, head and neck tumors and ovarian tumors, for which the SH3 domain of Grb-2 would be targeted. For example, tumors with associated amplification of receptors for EGF and PDGF could be inhibited by blocking activation of the Ras pathway through inhibition of the interaction between Grb-2 (SH3)and Ras. Furthermore, since the SH3 domain of Src family kinases are believed to be involved in activation of T-cells, B-cells, mast cells,

and NK cells and since the SH3 domains of the tyrosine kinases Tsk and Btk are believed to be involved in T-cell (Tsk SH3) and B-cell (Btk SH3) function an SH3 binding or blocking agent of this invention could be administered to a patient in need thereof to suppress immune function.

An SH3 binding or blocking agent identified by the method of this invention can be formulated into a pharmaceutical composition containing a pharmaceutically acceptable carrier and/or other excipient(s) using conventional materials and means. Such a composition can be administered to an animal, either human or non-human, for therapy of a disease or condition resulting from cellular events involving an SH3-mediated protein-protein interaction. Administration of such composition may be by any conventional route (parenteral, oral, inhalation, and the like) using appropriate formulations as are well known in this art. The SH3 binding or blocking agent of this invention can be employed in admixture with conventional excipients, ie, pharmaceutically acceptable organic or inorganic carrier substances suitable for parenteral administration.

Brief Description of the Figures

Figure 1 (SEQ ID NOS: 1 to 64) shows the amino acid sequences (SEQ ID NOS: 3 to 64) of peptides identified by screening a phage display library, designated the X_6PPIP library (SEQ ID NO: 3) (containing the sequences XXXXXXPPIP, where X is any amino acid), with Src, Fyn, Lyn, PI3K, Abl, Tsk and Yes GST-SH3 fusion proteins and with GRB2 (which contains two SH3 domains). The peptides displayed on the surface of M13 bacteriophage were inserted into the amino terminus of M13 gene III sequences (two residues from the amino terminus) such that all displayed peptides had the gene III residues alanine-glutamic acid at their N-terminus followed by the rest of gene III sequences. Amino acids that are in italics are geneIII residues. Isolates selected from the X_6PPIPG library are listed. The amino acids that are underlined were inserted into geneIII, adjacent to the six randomized residues. Amino acid preferences are highlighted in bold type. Numbers in parenthesis denote isolates that have the same nucleotide sequence. The isolate from the first library (containing the glycine to arginine change) is also shown.

Figure 2 (SEQ ID NOS: 65 to 113) shows the amino acid sequences of peptides identified by screening phage display libraries, RSLRPLXXXXXX (SEQ ID NOS: 65 to 99), PPPYPPXXXXXX (SEQ ID NOS: 100 to 108) and RLYRPLXXXXXX (SEQ ID NOS: 109 to 113) with various SH3 domains.

Figure 3. **Binding of phage-displayed peptides to Src and Abl SH3 domains.** Binding of phage-displayed peptides was measured using BIAcore® technology as described below. Interaction between the receptor (immobilized SH3 domain) and ligand

(phage-displayed peptide) is indicated by an increase in RU value. Traces have been corrected for bulk refractive index contributions of the unbound phage. **Upper right panel:** Binding of "RSSLRPLPPIP phage" to immobilized GST-Src SH3. **Upper left panel:** Binding of "RSSLRPLPPIP phage" to immobilized GST-Abl SH3. **Lower left panel:** Binding of "PPPYPPPPIP phage" to immobilized GST-Src SH3. **Lower right panel:** Binding of "PPPYPPPPIP phage" to immobilized GST-Abl SH3. These data illustrate the selectivity of binding of the phage-borne SSL sequences for the SH3 domains with which they were selected.

Figure 4 (SEQ ID NOS: 133 to 156) shows the amino acid sequences (SEQ ID NOS: 134 to 156) of peptides identified by screening a phage display library, designated the XXXXXRRLPPLPPP library (SEQ ID NO: 133), with GST-Src, Lyn and PI3K SH3 fusion proteins.

Figure 5 (SEQ ID NOS: 114 to 132) shows the amino acid sequences (SEQ ID NOS: 115 to 132) of peptides identified by screening a phage display library, designated the RSLRPLPPLPXXXXX library (SEQ ID NO: 114), with GST-Src, Lyn, Fyn, Yes and PI3K SH3 fusion proteins.

Figure 6 shows BIAcore® binding analysis of six phage displayed peptides to immobilized GST-Src or the GST-Lyn SH3 domains. Included in this analysis are the phage that display Src selected sequences (5) (SEQ ID NO: 149), Lyn selected sequences (6) (SEQ ID NO: 141), two phage which display core consensus sequences (1, 2) (SEQ ID NO: 8 and SEQ ID NO: 6), and two samples which did not have the arginine residue as part of the core consensus sequence and were correctly predicted to bind with lower affinity (3, 4) (SEQ ID NO: 39 and SEQ ID NO: 24).

Figure 7 shows the amino acid sequences of peptides identified by screening a phage display library, designated the GAAPPLPPRXXXXX library, with GST-Src, Lyn, Fyn, Yes and PI3K SH3 fusion proteins.

Figure 8 shows the amino acid sequences of peptides identified by screening a phage display library, designated the random 10mer library (XXXXXXXXXX), using NH2 terminal Crk, Lyn, Src and Fyn GST-SH3 fusion proteins.

In Figures 1-8 (and elsewhere), the three-letter symbols represent the following amino acids: ala = alanine; arg = arginine; asn = asparagine; asp = aspartic acid; asx = Asn and/or Asp; cys = cysteine; gln = glutamine; glu = glutamic acid; gly = glycine; his = histidine; ile = isoleucine; leu = leucine; lys = lysine; met = methionine; phe = phenylalanine; pro = proline; ser = serine; thr = threonine; trp = tryptophan; tyr = tyrosine; val = valine.

Figure 9. **Construction of GAL4-SrcSH3 DNA binding domain expression vector.** A segment of DNA encoding amino acids 84-145 of human Src SH3 is amplified by PCR using two primers, SH3PCR1 and SH3PCR2. The amplified DNA

is cleaved with restriction enzymes EcoRI and BamHI then ligated into the vector pSG424 (Sadowski and Ptashne, Nuc. Acids Res. (1989)18, 7539). Upon introduction into the appropriate cells, this vector produces a fusion protein consisting of the N-terminal portion of GAL4 (amino acids 1-147) fused to the Src SH3 domain. In the figure, the black box in front of the GAL4 portion represents the promoter region required for expression. Other vector elements, such as selectable markers, are not shown.

Figure 10. **Construction of SSL-VP16 activation domain vector.** Two synthetic oligonucleotides are annealed to form a double stranded DNA fragment encoding an SSL. The SSL DNA is ligated into the vector pCGN (Tanaka and Herr, Cell (1990) 60, 375-386). Upon introduction into the appropriate cells, this vector produces a fusion protein consisting of the SSL fused to the transcriptional activation domain of Herpes Simplex Virus VP16. In the figure, the black box in front of the SSL portion represents the promoter region required for expression. Other vector elements, such as selectable markers, are not shown.

Figure 11: **A system for identification of compounds that interfere with SH3 molecular interactions.** To screen for compounds that specifically interfere with SH3-mediated interactions, the two vectors from Fig 9 and Fig 10 are introduced into appropriate cells (by way of example, an not limitation, mammalian cell lines termed COS, CHO, Jurkat, Daudi or 293 cells) where both the GAL4-SH3 and SSL-VP16 fusion proteins will be produced. The cells contains a reporter gene under the control of GAL4. In the absence of competing molecules, the association of SH3 and the SSL will promote the stimulation of transcription through the GAL4 and VP16 partners. This results in expression of the reporter gene which can be detected by various methods as described in the text. Compounds that interfere with the SH3-SSL interaction will be detected as a reduced (or in some cases) increased level of reporter. Certain other reporter systems will promote cell growth in the absence (or in other cases in the presence) of an interfering compound. This system is not limited to mammalian cells; it may be readily adapted using appropriate expression vectors for use with yeast, insect cells, or bacteria. In addition, this sytstem is not limited to intact cells but may also be accomplished using cellular extracts or in vitro transcription/translation. Panel A: Representations of molecules in panels B-D. Panel B: A cell line expressing GAL4-SH3 and SSL-VP16. In the absence of a test compound, GAL4-SH3 binds to SSL-VP16 to promote expression of reporter gene. Panel C: In the same cell line as panel B, the presence of a test compound that binds to SH3 prevents association with SSL-VP16. The reporter gene is not expressed. Panel D: A control cell line expressing the GAL4-VP16 is used to test the ability of

the test compound to interact with other cellular components. The preferred SH3 inhibitory compound would not affect reporter gene expression in the control cell line.

Figure 12. **Alignment of 17 SH3 domains.** The SH3 domain regions from 17 proteins were aligned by the Megalign program (DNAstar, Madison WI) using the clustal method with a PAM250 weight table. Amino acids that are conserved in at least 10 out of the 17 proteins are shown above the sequences as Consensus #1. Two amino acids, W and G, designated by arrows, are found in all 17 SH3 proteins. Dashes represent no amino acid present were included where insertions in other SH3 domains occur. The *consensus homology domain*, is defined as the sequence from the first conserved amino acid (A) through the last conserved amino acid (V). All protein listed are from human except YSCH8263_5, a sequence from the yeast S. cerevisiae, that was included to demonstrate the homology between human and yeast SH3-like domains.

Figure 13. **Sequence Pair Distances for 17 SH3 Domains.** An analysis of percent similarity (and percent divergence) was tabulated using the Megalign program. The highest percent similarity between two SH3 domains is for Lck and Lyn (45.1%). The lowest percent similarity is between p40 and p85A (11%). These data were generated using a PAM250 residue weight table.

Figure 14 shows BIAcore® analysis of the ability of one SH3 ligand (SOS-derived peptide Ac-EVPVPPPVPPRRRPGGK-NH2; solid spheres •) and one SSL (Ac-SHRLYRPLPPLPGGK-NH2; hollow spheres o) to competitively inhibit the formation of SH3/ligand complex as described in Example 5. Assays were initiated by combining 1uM Grb-2(P204L) SH3 domain with various concentrations (100, 50, 25, 10, 5, 2.5, 1 or 0 uM) of peptides and pre-incubated at 25°C for >30 minutes. Complex formed between the SH3 domain and the immobilized ligand was measured using the BIAcore Biosensor® (Pharmacia Biosensor, Piscataway, NJ).

Detailed Description of the Invention

Identification of SH3 or SH3-like Domains

The term "SH3-like domain or a subdomain thereof" refers to a sequence which is substantially homologous to a Src homology region 3 (SH3 region), or a subdomain of an SH region preferably a conserved region of an SH region. The Src homology region is a noncatalytic domain of ~60 amino acids which was originally identified in the viral Fps and viral Src cytoplasmic tyrosine kinases by virtue of its effects on both catalytic activity and substrate phosphorylation (T. Pawson, Oncogene 3, 491 (1988) and I. Sadowski et al., Mol. Cell. Biol. 6, 4396 (1986)). SH3 domains have been found in a variety of eukaryotic proteins, some of which function in intracellular signal

transduction. Examples (including counterparts from various species) of SH3 domain-containing proteins include (1) members of the src-family protein tyrosine kinases (Src, Lyn, Fyn, Lck, Hck, Fgr, Yes), (2) Grb-2, which has two SH3 domains, (3) Sprk, a threonine/serine protein kinase, (4) Tsk, (5) Btk, (6) Txk and other Tec family members (see e.g. Haire et al, Hum Mol Genet 30, 897 (1994), (7) Vav, (8) GTPase Activating Protein (GAP), (9) p40, p47, and p67 proteins of the neutrophil oxidase complex, and (10) phosphatidylinositol 3' kinase, (11) Crk, (12) phospholipase C gamma, (13) Abl. SH3 domain-containing proteins have been identified in human, rodent, bovine, *C. elegans*, and yeast.

In **Figure 12**, primary amino acid sequences of a selection of 17 known SH3 domains have been aligned by the "clustal method" using multiple alignment software from DNAstar (Madison, WI). As shown in **Figure 12**, residues which are conserved in 10 out of 17 sequences are listed above the alignment as *Consensus #1*. This criteria identified 18 consensus amino acids and the region of each protein including these amino acids is know as the *consensus homology domain*.. The length of the consensus homology domain—that is, the number of amino acids from the first conserved amino acid of the domain (alanine, A) to the last conserved amino acid of the domain (valine, V) varies from 48 amino acids for Grb2 (C-terminal domain), p47 (C-terminal comain), p67 (both domains) to 65 amino acids for p85 (PI3'kinase). However, other SH3 consensus regions may be longer or shorter. Insertions or deletions compared to Src SH3 domain can occur anywhere within the consensus region. For example, the PI3K SH3 domain, neuronal Src SH3 domain and Sprk SH3 domains contain inserted amino acids between the VL and WW consensus motifs. All 17 of the SH3 domains shown in figure X contain a conserved tryptophan (W) and glycine (G) within the domain. Some SH3 domains, such as human Sprk SH3, contain all 18 consensus amino acids. Others contain 9-17 of the consensus amino acids. Such an analysis may be useful in identifying new SH3 or SH3-like domains from unknown DNA, RNA, or protein sequences. For example, a protein coding sequence can be aligned with the 18 consensus amino acids, or Src SH3 or any of the other SH3 or SH3-like domains and scored for the presence of the consensus amino acids. An SH3-like domain may be recognized by having between greater than 7 of these conserved amino acids but SH3-like domains may have fewer than 7 conserved amino acids. An alternative method for identification of SH3-like domains is through the use of sequence distance calculations such as shown in **Figure 13**. This figure shows the percent similarity between pairs of the 17 SH3 domain sequences in the alignment. The highest sequence identity among this set of SH3 domains was Lck SH3 compared to Lyn SH3 with a %sequence identity of 45.1. Two SH3-like domains may exhibit higher or lower values than shown here. The lowest value was 11.0% between p40 and p85 SH3 domains.

Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Certain SH3 or SH3-like domains may not match any of the 18 conserved amino acids nor exhibit significant homology with known SH3 domain sequences to be detected by computer alignment programs. These sequences may, nevertheless, exhibit the same or similar three-dimensional structure as known SH3 domains and function as an SH3-like domain. The three-dimensional structure of several known SH3 domains have been determined. SH3 domains are characterized as two anti-parallel beta sheets composed of 5 or 6 beta strands. Regions forming an alpha helix may or may not be present within the domain. SH3 or SH3-like domains may be recognized as having an SH3-like domain structure when solved by x-ray crystallography or NMR spectroscopy. Alternatively, a predicted structure by homology modeling may be used to identify a particular protein sequence as an SH3-like domain.

The identification of SH3 or SH3-like domains may be accomplished by screening a cDNA expression library with an SSL for known SH3 domains to isolate cDNAs for SH3 proteins. One could use PCR or low stringency screening with an SH3-specific probe. The SH3 domain or protein containing the SH3 domain may be isolated from naturally occuring sources (*e.g.* cells, tissues, organs, *etc*); produced recombinantly in bacteria, yeast or eukaryotic cells; produced *in vitro* using cell free translation systems; or produced synthetically (*e.g.* peptide synthesis).

The subject invention is relevant to SH3 and SH3-like domains as described in the foregoing paragraphs.

Further definitions

The following additional terms are defined in the interest of clarity and to aid in understanding the invention.

A *specific SH3 ligand* or *SSL* is a ligand which binds selectively to an SH3 domain of interest.

As used herein, "*selective*" binding to, blocking or inhibiting means "capable of binding to, blocking or inhibiting more effectively with respect to the SH3 domain of interest than with respect to at least one other SH3 domain". This selectivity may be manifested biochemically as an increased on-rate ($K_{association}$) and/or a decreased off-rate ($K_{disssociation}$) for the SH3 domain of interest, *e.g.* as determined by kinetic methods including BIAcore® analysis or as a lower Kd as measured by any conventional binding assay methodology or as a larger Gibbs free energy of binding

(G_0) as measured or calculated by standard methodologies. Alternatively, this selectivity may be manifested as a lower IC50 of a test substance for competitive inhibition of binding of an SH3 ligand to an SH3 domain of interest, whether determined *in vitro* or in cell-based assays (*e.g.*, via expression of a reporter gene) or a lower ED50 of a test substance in a cell, organ or animal model where 1) the involvement of an SH3 target in the outcome of the assay can be demonstrated; and 2) binding of the test substance to the SH3 target can be demonstrated *in vitro* by any of the above means. In particular cases, the degree of selectivity is preferably at least 2-fold, and even more preferably at least 3-fold, as measured by any of the foregoing types of measurement. Degrees of selectivity of greater than about 5-fold and greater than about 10-fold and higher are achievable in individual cases.

A substance identified by the methods of this invention, which selectively binds to an SH3 domain of interest, or blocks or inhibits protein-protein, protein-peptide, protein-nucleotide, protein-polynucleotide, protein-lipid, protein-carbohydrate or protein-small molecule interactions mediated by an SH3 domain of interest, is referred to as an SH3 binding agent or SH3 blocking agent. Substances to be assessed for their ability to bind selectively to an SH3 domain of interest can be obtained from a variety of sources, including for example microbial broths, cellular extracts, conditioned media from cells, synthetic compounds and combinatorial libraries. Once an agent has been identified as an SH3 binding or blocking agent, it can be produced using known methods, such as by recombinant methods of protein production or chemical synthesis. It can also be obtained from the source in which it was initially identified (*e.g.*, a microbial broth, cellular extract, *etc.* as above). Specific SH3 ligands identified by screening bacteriophage display libraries are themselves useful in certain applications as SH3 binding or blocking agents. Binding or blocking agents for possible therapeutic applications can be further evaluated by conventional methods with respect to toxicological and pharmacological activity.

As mentioned above, pharmaceutical agents which interfere with the formation or stability of signaling complexes (such as those formed by signaling proteins containing one or more SH3 domains with their natural ligands) can be used to treat or prevent the diseases or their pathological effects mediated by such complexes—but only if these interfering agents have sufficient binding selectivity. Unfortunately, until the work described herein, very little was known about the natural ligands or the binding specificities of the various SH3 domains to determine whether selective interference with SH3 ligand binding would be possible, and if it were, to permit the identification or design of agents capable of so acting. Indeed, the limited apparent sequence requirements (proline rich sequences) for binding to SH3 domains, together

with the limited expected affinities of interactions therewith were considered discouraging, especially in comparison to emerging information for SH2 domains.

Despite such considerations, the present inventors have now developed a method for identifying ligands, referred to as specific SH3 ligands (SSLs), which bind selectively and specifically to SH3 domains of various proteins with high affinity. These ligands are useful in the subject method for identifying agents which selectively and specifically block or inhibit protein-protein interactions via a preselected SH3 domain. Such agents may be useful in the therapy of diseases and conditions mediated by such protein-protein interactions. The specific SH3 ligands described herein, such as those presented in Figures 1 to 8, are also SH3 binding agents.

The following is a description of our work, including the discovery of the binding specificity of SH3 domains, and applications of these important findings.

Identification and characterization of SSLs

SSLs are preferably selected from combinatorial arrays (biased or unbiased) of candidates as can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form or as phage display libraries. Combinatorial libraries can also be synthesized of peptides containing one or more D-amino acids and/or non-naturally occuring amino acids. Libraries can also be synthesized of peptoids and non-peptide synthetic moieties. Such approaches provide access to candidate SSLs containing novel peptide sequences, peptide sequences not found in nature and novel non-peptide strucutres.

SSLs are selected from the candidate molecules using the SH3 domain of interest, typically in the form of the protein containing the SH3 domain, a fragment of such protein, a peptide or fusion protein containing the amino acid sequence of the SH3 domain of interest. Additional guidance in this regard is provided below.

Selection may be effected by combining one or more candidate SSLs with the SH3 domain of interest, incubating the resultant mixture under conditions permitting the formation of an SH3 domain-ligand complex, separating the complexes so formed from uncomplexed candidates and identifying the candidates which bound to the SH3 domain.

We have found phage display to be particularly effective in the identification of peptide SSLs. Interestingly, SSLs so identified usually comprise novel peptide sequences not represented by known peptide or DNA sequences. We have developed this approach for identifying SSLs of varying lengths ranging from minimal core sequences to longer sequences which also contain one or more, and often preferably several more, flanking amino acid residues, as discussed in further detail below. See also Rickles et al, 1994, EMBO Journal 13(23):5598-5604, the full contents of which

are incorporated herein by reference. By the means disclosed herein, SSLs may be obtained with measurable K_D values, with respect to an SH3 domain of interest, in the range of about 7 - 30μM, more preferably in the range of about 2 - 8 μM, and even more preferably in the range of about 0.22 μM to about 2 μM. In some cases SSLs with even lower K_D values for an SH3 of interest, which may be obtained as described herein, will be preferred. K_D values may be readily measured, for example, using BIAcore techniques, as disclosed herein, or by flourescence spectroscopy (see e.g. Chen et al, J Am Chem Soc 115, 12591-12592 (1993)).

Preparation of SH3-like Domain for Selection of SSLs or Compound Screening

In one embodiment, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts of from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent a completely degenerate or biased array. By selection against the SH3 domain of interest one selects phage bearing inserts which bind to the SH3. This process can be repeated through several cycles, typically 2 to about ten, of reselection of phage that bound. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis is conducted to identify the sequence(s) of the expressed peptides. By analysis of the peptide sequence of the selected phage a preferred or consensus SH3 ligand sequence can be identified. The minimal linear portion of that sequence that binds to the SH3 domain of interest is termed the "core". This can be determined using synthetic peptides, with phage display or other techniques. It will often be preferred to start with phage inserts of seven to ten residues and identify a consensus selected sequence. One can repeat this procedure using a biased library containing inserts containing part or all of the consensus sequence plus one or more additional degenerate residues upstream or downstream thereof. Sequential rounds of such selections permit identification of SSLs containing flanking sequence. (Identification of core and flanking sequences may also be conducted by analgous conventional procedures with synthetic peptide libraries in place of phage libraries. In that case, the design of the peptide library is analogous to the design of the phage inserts.)

SH3 expression vectors can be constructed by ligating into a conventional expression vector the DNA sequence encoding the consensus homology domain for the SH3 domain of interest, alone or preferably with additional flanking sequence. For example, for human Src the SH3 consensus homology domain includes amino acids 91-140. We have prepared SH3 domain protein from E. coli expression vectors using amino acids 84-145, which contain an additional 7 amino acids on the N-terminal side of the homology domain and 5 amino acids on the C-terminal side. These additional amino acids may provide stability to the SH3-like domain, may improve

its production in the expression cells, may improve its ability to interact with ligands or other proteins or be necessary for linking to a fusion protein for reasons discussed below.

The SH3 domain may be expressed within all or part of its natural context, as an isolated SH3-like domain, in a tandem array of SH3-like domains containing two or more of the same or different domain, or as a fusion protein with other unrelated domains including but not limited to SH2-like domains, protein kinase domains glutathinone S-transferase (GST), epitope tags, kinase recognition sequences, maltose binding protein, signal sequences, biotin-modification sequences.

The SH3-like domains may be modified to:

- facilitate purification *e.g.* by expression as a fusion to glutathione-S-transferase, maltose binding protein, metal-chelation sequences (poly-histidine), protein A or others;
- facilitate identification or quantitation, *e.g.* by covalent modification using biotin, fluorophores, chromophores, scintillons, spin labels, radioactive or non-radioactive isotope tags, magnetic particles, metal coloids, *etc;*
- generate a measurable signal, *e.g.* by expression as a fusion to alkaline phosphatase, horse radish peroxidase, green protein, luciferase, *etc;*
- adhere to defined solid supports, *e.g.* by expression as a fusion to an epitope tag or other antigenic domain; engineered to provide unique or uniquely accessible protein features *e.g.* N-terminal serine, cysteine, lysine or others, *etc.*
- remove undesirable features that do not effect SH3 function but pose experimental complications, *e.g.* by mutation of cysteines that participate in unnatural domain dimerization
- improve stability under conditions of binding assays (*e.g.* by altering the natural coding sequence to encode cysteines that form stabilizing disulfides).

Selection of SSLs

Initially, a phage library containing displayed peptides consisting of six randomized residues ("X" represents any amino acid) bracketed by glycines (-GXXXXXXG-) (SEQ ID NO: 157), was screened using the Src SH3 domain, prepared as part of a GST fusion. Phage which bound the GST-SH3 fusion were isolated and sequenced. Enrichment was observed when using GST-Src SH3 but not when using GST. The sequencing of ten Src SH3 domain-selected phage showed that each phage insert had the same nucleotide sequence, which encoded a proline-rich insert. Furthermore, the Src SH3 selected peptide was seven residues in length, not six, due to a point mutation that converted a glycine residue to arginine:

(SEQ ID NO: 157) GXXXXXG (starting library)
(SEQ ID NO: 158) RSLPPIPG (Src SH3-selected sequence)

This fortuitous mutation probably occurred during phage propagation. That this was the only peptide selected when using the src SH3 domain suggests that a Src SH3 ligand should be, at a minimum, seven residues in length, proline-rich and contain an arginine residue as part of the recognition sequence.

A second and *biased* library was prepared, in which the phage displayed peptides containing six random amino acid residues flanking the tetrapeptide proline-rich sequence identified in the first screen (-X_6PPIPG-). This new library, designated the X_6PPIP library (SEQ ID NO: 3), was screened using GST- Src SH3. The amino acid sequences of peptides identified in this manner are shown in figure 1. These results support the view that the *Src* SH3 recognition sequence is proline-rich and at least seven amino acid residues in length. Of the Src SH3-selected sequences, while no amino acid preferences were observed at positions 1, 2 and 3 of the six randomized residues, arginine was preferred at position 4, proline at position 5, and leucine at position 6. That arginine is preferred at the fourth position supports our view that the point mutation observed in the isolate from the first library (changing glycine to arginine) was critical for Src SH3 selection.

The X_6PPIPG library was also screened using Fyn, Lyn, Yes, PI3K, Tsk, Abl GST-SH3 fusion proteins and with GRB2 (a 25 kD protein containing two SH3 domains). Each of these proteins enriched phage from the X_6PPIPG library. The results from sequencing selected phage are shown in Figure 1. Like the Src SH3 domain, the Fyn, Lyn, Yes and PI3K SH3 domains and GRB2 prefer ligands that contain arginine at position 4, proline at position 5 and leucine at position 6. However, the Lyn, Yes, PI3K and GRB2 SH3 domains have additional binding preferences.

In contrast to the ligands identified using the Src, Fyn, Lyn, Yes, PI3K and GRB2 SH3 domains, the Tsk and Abl SH3 domains selected ligands that were quite disparate. The Abl SH3 domain, for example, showed preferences for peptides containing prolines at the first three positions of the six randomized residues. Furthermore, tyrosine and proline were selected at the 4th and 6th position respectively.

To investigate the sequence preferences in the "PPIP" region of the SH3 ligands, three new biased libraries were generated. The library RSLRPLX6 was prepared to identify additional Src, Fyn, Lyn and PI3K SH3 ligands. The library PPPYPPX6 was prepared and screened to further define Abl SH3 binding preferences and the library RLYRPLX6 was used to evaluate GRB2 ligand preferences.

Bacteriophage were selected from these new libraries and sequenced. As shown in Figure 2, the proline motif found originally by screening the GX_6G library using the Src SH3 domain was again selected, this time by all SH3 domains whose binding preferences were analyzed. The C-terminal consensus sequences derived from the $RSLPRLX_6$ library are PPLPXP (Src), PP(I/L)PXX (Fyn), PPLPXP (Lyn), PPLPPP (PI3K). When using the PPPYPPX6 library, the Abl SH3 domain prefers peptides that contain the C-terminal residues PP(I/V)PXX. GRB2 selects SH3 ligands that contain the residues PPLPXX. The minimum number of selected consensus residues that suffices for recognition by an SH3 domain we here term the core recognition sequence. The Src SH3 domain, for example recognizes and prefers the sequence RPLPPLPXP. In contrast, the Abl SH3 domain prefers the core consensus sequence PPPYPPPP(I/V)P.

To rapidly evaluate the binding properties of the phage-displayed ligands, and to gain insight into SH3 domain selectivity, the relative binding affinity of phage isolates was measured by a novel application of surface plasmon resonance. The binding of phage-displayed ligands to immobilized GST-SH3 proteins was analyzed using a BIAcore Biosensor®. To obtain preliminary evidence whether the Abl and Src selected phage showed selective binding for each of these SH3 domains, relative binding affinity was measured. As shown in Figure 3, binding of phage containing the Src SH3 selected consensus sequence was readily detected using GST-Src SH3 but not using GST-Abl SH3. Conversely, phage displaying the Abl SH3-selected sequence PPPYPPPPIP bound to GST-Abl SH3 but binding to GST-Src SH3 was not observed.

Peptides corresponding to phage displayed sequences may be synthesized and used to more accurately quantify their binding to SH3 domains, *e.g.* using surface plasmon resonance techniques. For example, using the "Src" and "Abl" peptides, binding was studied *in vitro* by surface plasmon resonance. Again, Src and Abl SH3 binding was selective. The Abl SH3 domain bound with high affinity to the Abl and not the Src-selected peptide. Conversely, the human and avian Src SH3 domains recognized the Src but not the Abl-selected sequence. The results are summarized in Table 1.

Table 1

	Abl (X ± SEM)	hSrc (X ± SEM)	aSrc (X ± SEM)
aePPPYPPPIPggk	2.0 ± 0.2	NDB	NDB
aeRSSRPLPPIPggk	NDB	19.5 ± 1.1	58.5 ± 10
kggAPTMPPPLPPVPPg	66.3 ± 19.3	NDB	NDB

Table 1. *Binding affinities of Src and Abl SH3 domains for peptides.* Binding of human Src (**hSrc**), avian Src (**aSrc**) and abl (**abl**) SH3 domains to Src-selected (**aeRSSRPLPPIPggk**), abl-selected (**aePPPYPPPPIPggk**) and 3BP1 (**kggAPTMPPPLPPVPPg**) sequences. K_D values are umolar and represent the Mean ± SEM (n=3). **NDB**: no detectable binding at a protein concentration of 50 μM.

We have exploited these methods using biased phage libraries to more fully define SH3 domain ligand binding preferences. We have discovered that SH3 domains that prefer almost identical core sequences may still have quite distinct binding preferences. Additional biased libraries were prepared which displayed the core sequence RPLPPLPPP adjacent to five random upstream residues (X_5RPLPPLPPP). This library was screened using the Src, Fyn, Lyn, Yes and PI3K SH3 domains. As shown in Figure 4, selected ligands had preferences at positions upstream of the core sequence. Some observed preferences are shared by subclasses of SH3 domains while other preferences are distinct for a particular SH3 domain. Additional sequence preferences were observed when a library was screened which contained five random residues downstream of core sequences (Figure 5).

We have analyzed the effect that flanking sequences can have on binding affinity. As shown in Figure 6, phage displaying a core sequence binds with similar affinity to the Src and Lyn SH3 domains as determined by BIAcore® analysis. In contrast, phage displaying an SH3 ligand that consists of the core sequence and Src SH3-selected upstream flanking sequences bind with higher affinity to the Src SH3 domain than to the Lyn SH3 domain. Conversely, phage displaying a ligand consisting of the core sequence and Lyn SH3-selected upstream flanking sequences bind with higher affinity to the Lyn SH3 domain than to the Src SH3 domain. Binding analysis, using peptides instead of phage reveals that a ligand consisting of the Src-selected upstream flanking sequence and the core sequence (VSLARRPLPPLPPP, the

core recognition sequence is underlined) binds with 33-fold higher affinity (Kd= 0.8 um) to the Src SH3 domain than peptides that contain only the core sequence (GGSGGRPLPPLPPP or RPLPPLPP, Kd's = 27um). The results just described demonstrate that sequences which flank the core recognition sequence are important determinants of both binding affinity and binding selectivity.

The solution structures of Src SH3 domain-ligand complexes revealed that the Src SH3 domain can bind the Src SH3 preferred core ligand sequence (RPLPPLP) when it is in the opposite orientation (Feng et al (1994), Science 266, 1241-1247). With class II ligands, the conserved core arginine residue is at the COOH-terminus, not the amino-terminus (RPLPPLP) as is found for what is now termed the class I Src SH3 ligand core sequence. We designed a phage library that displayed a class II core and five random downstream residues (GAAPPLPPRX_5, the conserved arginine residue is in bold type, the core recognition sequence is underlined) to determine whether there might be preferred flanking sequences for class II ligands. The library was screened using the Src, Fyn, Lyn, Yes and PI3K SH3 domains. The results from sequencing selected phage is shown in Figure 7. As with class I ligands, there are SH3-specific sequence preferences adjacent to the class II ligand core. Furthermore, we find that for each SH3 domain analyzed, the sequence preferences adjacent to class I and class II cores are distinct. For example, the Src SH3 domain selects the residues XSLXX adjacent to the arginine of the class I core (XSLXXRRPLPPLPPP, the class I core sequence is underlined, the critical arginine of the class I core is in bold type). In contrast, Src SH3 ligands identified using class II libraries have the consensus sequence NRPRL adjacent to the class II arginine residue (GAAPPLPPRNRPRL). Binding analysis, using peptides, reveals that a ligand consisting of a class II core and Src-selected downstream flanking sequences (GAAPPLPPRNRPRL, the core recognition sequence is underlined) binds with fifty-fold higher affinity (Kd= 0.22 μm) to the Src SH3 domain than peptides that contain only the core sequence (GAAPPLPPR, Kd = 11.4μm).

The Src SH3 selected flanking sequences are important determinants of Src SH3 domain recognition. To determine how specific these residues are for SH3 domain recognition, binding assays were performed using the Lyn, PI3K, Fyn and Yes SH3 domains As shown in Table 2, a peptide containing class 1 core residues and the Src SH3 selected flanking sequence (VSLAR) binds at least 50 fold better to the Fyn and Yes SH3 domains compared to a peptide containing just the core recognition sequence. In contrast, the Src SH3 selected flanking sequences only marginally affect (2x) the affinity of the interaction of the core ligand residues with the PI3K SH3 domain and have no effect on binding to the Lyn SH3 domain.

Table 2

Class I SH3 Ligands

Core: RPLPPLPGGK
Core+Src SH3 selected flanking sequence: VSLARRPLPPLPGGK

Binding assay (Kd: μM)

SH3 Domain	Core peptide	Core+Src flank
Src	27.0	0.8
Lyn	5.4	4.0
PI3K	30.8	15.9
Fyn	29.1	0.7
Yes	62.0	0.7

Class II SH3 Ligands

Core: KGGGAAPPLPPR
Core+Src SH3 selected flanking sequence: KGGGAAPPLPPRNRPRL

Binding assay (Kd: μM)

SH3 Domain	Core peptide	Core+Src flank
Src	11.4	0.22
Lyn	27.6	6.2
PI3K	n.d.	n.d.
Fyn	99.5	1.7
Yes	>99.5	3.6

(n.d.= not determined)

Table 2. *Binding affinities of the Src, Fyn, Lyn, Yes and PI3K SH3 domains to class 1 and class 2 core peptides and core peptides containing Src SH3 selected flanking sequences.* Kd values were determined by Tryptophan Fluorescence measurements as described (Chen et al (1994) J. Am. Chem. Soc. 115, 12591-12592)

Similar results to those described above were obtained when comparing a class 2 core peptide with one that also has the C-terminal residues (NRPRL), identified using the Src SH3 domain. As observed with the Src SH3 domain, the C-terminal flanking sequences dramatically increase the affinity of core ligand binding to the Fyn and Yes SH3 domains. In contrast, binding to the Lyn SH3 domain was only marginally (about 4 fold) increased.

Binding was also analyzed using the Src SH3 domain and class 2 peptides that contained PI3K, Lyn or Fyn selected flanking sequences. As discussed above, the Src SH3 domain binds to a class 2 core (KGGGAAPPLPPR, the core recognition sequence is underlined) with a Kd = 11.4 μM. The addition of the Src SH3 selected C-terminal residues NRPRL increases the affinity about 50 fold (Kd = 0.22 μM). A peptide which consists of a class 2 core and Fyn selected sequences (NPRPL) binds to the Src SH3 domain with a Kd of 1.8 μM (6 fold increase in affinity). In contrast, class 2 core peptides containing Lyn (PSWMS) or PI3K (PPRPA) selected C-terminal residues bind with affinities similar to that observed with a class 2 core peptide (7.4 μM and 23 μM respectively).

Using the methods described above and elsewhere herein, one may thus identify peptides (and non-peptide agents) which selectively bind to an SH3 domain-containing protein of interest. SH3 core ligand sequences may be distinct (for example, see Figure 3; Src SH3 and Abl SH3-selected core sequences). In addition, residues that flank the core sequence can be distinct (for example, see Figures 4, 5 and 7). These flanking residues are important for selective recognition by SH3 domains. Such SSLs may then be used to identify agents which selectively block or interfere with the binding of a selected SH3 domain containing protein to its ligands without substantially interfering with the normal binding of other SH3 domain containing proteins, whether or not within the same family.

Once an SH3 binding agent has been identified based on its ability to prevent binding of the specific SH3 ligand to the SH3 domain of interest, it can be further characterized, after it is separated, if necessary, from the test mixture. For example, the affinity with which the SH3 binding agent binds the SH3 domain of interest (or other SH3 domains) or otherwise inhibits SH3-mediated interactions, its composition or chemical nature and its biological/pharmacological activity can be assessed, using known methods. Those SH3 binding agents which do not bind to or interefere with other SH3 domains, or which do so with very low affinity, are of particular interest, since they will provide a basis for pharmaceutical agents useful as highly specific inhibitors of an SH3 domain-containing protein of interest which do not inhibit other SH3 domain-containing proteins.

In vitro assays

SH3 binding or blocking agents may be identified based on their ability to inhibit formation of an SH3/ligand complex or to bind to SH3 domains using BIAcore® and allied technologies (See Example 5) or based on their ability to inhibit formation of an SH3/ligand complex using solid phase binding assays.

SH3 ligand domains

SH3 ligand domains for use in *in vitro* assay will generally contain at least about 6-10 amino acid residues, and in some cases up to about 100 amino acid residues or more, although generally they will contain between 6 and about 40 amino acid residues. Ligand domains may be selected from known ligand sequences for an SH3 domain of interest, may be identified by the phage display method disclosed in Example 1, or otherwise identified from cell broth, natural extracts, synthetic libraries, *etc.* Ligands may also be phage-displayed peptides, peptide "side chains" displayed on non-peptide backbones (peptoids), cells or cell fragments, organic molecules, *etc.* or hybrids of the above. The SH3 ligand domains may be modified to:

- facilitate purification *e.g.* expression as a fusion to glutathione-S-transferase, maltose binding protein, metal-chelation sequences (poly-histidine), protein A or others;
- facilitate identification or quantitation, *e.g.* covalent modification using biotin, fluorophores, chromophores, scintillons, spin labels, radioactive or non-radioactive isotope tags, magnetic particles, metal coloids, *etc;*
- generate a measureable signal, *e.g.* expression as a fusion to alkaline phosphatase, peroxidase, green protein, luciferase, *etc;*
- adhere to defined solid supports, *e.g.* fusion to an epitope tag or other antigenic domains; engineered to provide unique or uniquely accessible features *e.g.* N-terminal serine, cysteine, lysine or others, *etc.*

Alternatively, SH3 ligand domains may be synthesized on solid supports (synthetic resins; glass, silica, metal; magnetic beads; others) and assayed without detachment.

Labelling with fluorochromes, reactive groups or biotin can be accomplished using known methods (for example see Chemistry of Protein Conjugation and Cross-Linking, 340 pp, S. S. Wong, CRC Press, Inc., Boca Raton, FL, 1991). Examples of suitable fluorochomes include fluorescein and derivatives, Rhodamine B and derivatives, fluorescamine, eosin isothiocyanate, dansyl chloride, 7-amino-4-methylcoumarin, N-(4-anilino-1-napthyl)maleimide, 4'6-diamidino-2-phenylinole; examples of suitable reactive groups include diethylenetriaminepentaacetic anhydride, N-ethylmaleimide, 2-iminothiolane HCl; biotinylation reagents include

iodoacetyl-biotin, 1-biotinamido,-4-maleimidomethyl-butane and N-hydroxysuccinimidobiotin.

Preparation of radiolabelled materials can be accomplished using, for example, kinase catalysed 32P phosphorylation (see e.g. *Protein Phosphorylation, Part A.* (1991), T. Hunter and B. M. Sefton, eds., Methods in Enzymology Vol. 200, Academic Press, Inc, New York, NY, 763 pp.), chloramine-T (Greenwood et al, Biochem. J. 89:114, 1963), lactoperoxidase (Marchalonis et al, Biochem. J. 124:921, 1971), Bolton-Hunter reagent (Bolton and Hunter, Biochem. J.133:529, 1973) or Iodo-beads mediated 125I-iodination (Markwell Anal. Biochem. 125:427, 1982) or tritium incorporation by reductive methylation (Tack et al., J. Biol. Chem. 255:8842, 1980).

In vivo assays

Among other features, this invention provides genetically engineered cells, preferably eucaryotic cells, useful for high throughput identification of compounds which interfere with, inhibit or otherwise impede interactions mediated by an SH3 domain. The cells of this invention contain and are capable of expressing a reporter gene and genes encoding two fusion proteins. The two fusion proteins are capable of binding to each other and when they do so, of detectably activating the transcription of the reporter gene.

The first fusion protein is a DNA-binding fusion protein which contains a first domain capable of binding to a regulatory element linked to the reporter gene and a second domain which comprises one member of an SH3 binding pair.

The second fusion protein is a transcriptional activating fusion protein which contains a first domain capable of activating the transcription of a gene with which it becomes associated and a second domain which comprises the other member of an SH3 binding pair.

An SH3 binding pair consists of a first peptide sequence comprising an SH3 domain of interest and a second peptide sequence comprising a specific ligand for that SH3 domain of interest. Viewed from this perspective, one of the fusion proteins contains an SH3 domain, the other contains a ligand for that SH3 domain. The two fusion proteins contain additional domains capable, upon complexation of the fusion proteins, of activating transcription of a reporter gene under the expression control of a DNA element responsive to those additional domains.

One or both of the fusion proteins may also contain one or more optional domains or elements, including a domain capable of regulatably sequestering the fusion protein at a desired cellular location or compartment.

In the practice of this invention, cells containing the components mentioned above are cultured under suitable conditions and in a suitable culture medium

permitting cell growth and a detectable level of expression of the reporter gene. A test substance is added to the medium, the cells are cultured for an appropriate incubation period and the level of expression of the reporter gene is measured. A decrease in the level of expression of the reporter gene indicates that the test substance is a candidate inhibitor of the binding of the SH3 binding pair.

The assay may be repeated using a control cell containing and capable of expressing the reporter gene and its associated control elements and a gene encoding a control fusion protein. The control fusion protein contains both the DNA binding domain and the transcriptional activating domain which are separately present in the respective DNA-binding and transcriptional activating fusion proteins, previously described. Alternatively, two fusion proteins may be used, each containing one half of a binding partner with different binding specificities then those found for SH3 domains and the protein-protein interaction assayed as described above. The failure of the test substance to significantly alter the level of expression of the reporter gene confirms its activity as an SH3 binding inhibitor.

in vivo Components & their use

1 Host cells

Any cells, preferably eucaryotic cells, which can be transformed with heterologous DNA and which can be grown or maintained in culture may be used in the practice of this invention, including yeast and mammalian cells. Mammalian cells include those of mouse, hamster, rat, rabbit, dog, cow or primate, including human, origin. They may be of a wide variety of tissue types, including mast cells, fibroblasts, osteoclasts, osteoblasts, T cells, macrophages and neutrophils, and may be primary cells or cell lines.In one embodiment, cell-free transcription systems prepared from cells as described above or prepared from purified components may be used in lieu of cells.

As discussed below, the cells are engineered by the introduction of heterologous DNA(s) encoding (a) the DNA-binding fusion protein, (b) the transcriptional activating fusion protein and (c) a reporter gene under the transcriptional control of a promoter responsive to the complex of both fusion proteins. The DNA molecule encoding those components may be introduced into the cells as separate DNAs or linked together.

2 Reporter Gene Constructs

Suitable reporter genes include any gene whose expression can be detected. Such genes include those encoding an enzyme which catalyzes a reaction resulting in a

detectable color change, an enzyme that permits continued growth in the presence of an otherwise toxic agent, a repressor or suppressor of a biological function which is activated or enhanced by the reduced expression of the "reporter" gene. Examples of reporter genes which may be used in the practice of this invention include genes encoding enzymes such as luciferase, beta-galactosidase, alkaline phosphatase or horseradish peroxidase which catalyze reactions generating readily detectable products. Alternatively, the reporter gene may encode a readily detectable cell surface marker such as Thy1. One well known reporter gene is the gene encoding secreted alkaline phosphatase (SEAP). The production of SEAP may be monitored from the culture media without the need to lyse the cells. This permits convenient detection, using a colorometric or fluorescent assay, at different time points after the addition of a test compound to the cell culture.

The reporter gene will be linked to a promoter which is responsive to a transcriptional activater capable of binding to the promoter and activating transcription therefrom. The promoter may be a synthetic promoter consisting of multiple binding sites, upstream of a TATA box element, for the DNA binding domain of a transcription factor. The number of such binding sites can be adjusted to optimize the overall level of reporter gene expression induced by the fusion proteins of this invention. Examples of such binding sites include GAL4 binding sites (see *e.g.* Acheson, 1980, DNA Tumor Viruses, (Cold Spring Harbor, NY, Tooze, J. ed.) pp 151-160; Peden et al, 1980, Science 209:1392-1396; Liu & Green, 1990, Cell 61:1217-1224; and Ptashne & Gann, 1990, Nature 346:329-331) for embodiments using GAL4 as the DNA-binding domain of the DNA-binding fusion protein and VP16 as the transcriptional activating domain of the cognate transcriptional activating protein. Additional guidance in the use of DNA-binding and transcriptional activating domains and promoters responsive thereto may be obtained from Fields et al, US Patent No 5,283,173 (1 Feb 1994) and Vasavada et al, 1991, Proc Natl Acad Sci USA 88:10686-10690.

3 Fusion protein Constructs

The two fusion proteins of this invention will typically contain either (a) one or more DNA binding domains or (b) one or more transcriptional activating domains, in addition to one or more SH3 domains or ligands thereto. One embodiment of the invention involves a pair of fusion proteins, the first containing one or more DNA binding domains and one or more SH3 domains, the second containing one or more transcriptional activation domains and one or more SH3 ligand domains. Another embodiment involves a pair of fusion proteins, wherein the first contains one or more DNA binding domains and one or more SH3 ligand domains, the second containing

one or more transcriptional activation domains and one or more SH3 domains. These are depicted schematically in Figures 9-11. The DNA binding and transcriptional activation domains may be derived from the same or different transcription factors.

Fusion proteins of this invention may also contain a targeting sequence providing for translocation of the protein to the nucleus. For example, such a targeting sequence may have a plurality of basic amino acids, referred to as a bipartite basic repeat (reviewed in Garcia-Bustos et al, Biochimica et Biophysica Acta (1991) 1071, 83-101). This sequence can appear in any portion of the molecule internal or proximal to the N- or C-terminus and results in the fusion protein being inside the nucleus.

The transcription components can be endogenous or exogenous to the cellular host. If the transcription factors are exogenous, but functional within the host and can cooperate with the endogenous RNA polymerase (rather than requiring an exogenous RNA polymerase, for which a gene could be introduced), then an exogenous promoter element functional with the fused transcription factors can be provided with a second construct for regulating transcription of the target gene. By this means the initiation of transcription can be restricted to the gene(s) associated with the exogenous promoter region, i.e., the target gene(s).

A large number of transcription factors are known which require two subunits for activity. Alternatively, in cases where a single transcription factor can be divided into two separate functional domains (*e.g.* a transcriptional activator domain and a DNA-binding domain), so that each domain is inactive by itself, but when brought together in close proximity, transcriptional activity is restored. Transcription factors which can be used include yeast GAL4, which can be divided into two domains as described by Ma and Ptashne (Cell (1987) 48, 847-853) and Fields and Song, supra. One may use, for instance, a fusion of GAL4(1-147)-SNF1 and SNF4-GAL4(768-881), where the SNF1 and -4 may be replaced by the subject binding proteins as binding domains. Combinations of GAL4 and VP16 or HNF-1 and VP16 can be employed. Each of these proteins will have a linked partner, part of an interacting pair. The interaction (binding partner recognition) will bring the transcriptional activation domain (VP16) to DNA (GAL4 or HNF-1) and transcription will be induced (Figure 11). Other transcription factors that could be used for these studies include , but are not restricted to, the members of the Jun, Fos, and ATF/CREB families, Oct1, Sp1, HNF-3, the steriod receptor superfamily, and the like.

DNA binding domains

Preferred DNA-binding domains for use in this invention meet the following criteria. First, they should be small, stably folded domains. Second, there should be background information available on the nature of the interaction of the domain with

DNA and with other proteins. Third, the preferred domains should bind DNA as monomers, although domains that bind as dimers could be accomodated. Fourth, they should bind DNA with relatively rapid kinetics, particularly with regard to the association rate. Examples of DNA-binding domains that meet some or all of these criteria are domains of the homeodomain class, the zinc-finger class, and the paired-box class, for which numerous examples are known in the literature.

In cases where detailed information on the molecular contacts of the protein with DNA is available, the DNA recognition specificity of the protein may be engineered by amino acid substitutions.

A preferred feature for DNA binding domains is the lack of interactions with other cellular proteins. Alternatively, the nature of such interactions should be known with precision, so that these interactions could be abrogated by suitable amino acid substitutions within the DNA-binding domain. In addition, the DNA binding domain may be unique for the cell type utilized. The DNA binding element, when introduced into cells, will preferably not be occupied in the absence of exogenously added DNA binding domain. The extent of occupancy of the DNA binding element may be determined by measuring the level of gene transcription when the DNA binding domain is part of a reporter gene's regulatory elements, in the absence and presence of exogenously added cognate transcription factor.

Cell-type specific factors (not found in the cells used for the *in vivo* assay), DNA binding domains with altered binding specificity and procaryotic DNA binding domains are well suited for the methods described. One suitable DNA-binding domain is derived from the Phox1 protein (Grueneberg, D. A., Natesan, S., Alexandre, C. and Gilman, M. Z. (1992). Human and Drosophila homeodomain protein the enhance the DNA-binding activity of serum response factor. Science 257, 1089-1095; Genbank accession number: M95929). This protein is a member of the homeodomain class. A 69 amino-acid domain derived from Phox1 is sufficient to bind DNA. High-affinity DNA recognition sequences for the protein have also been identified, as have amino acid substitutions that change its DNA recognition specificity and that affect its ability to interact with certain endogenous proteins in human cells. (Grueneberg et al)

A second suitable DNA-binding domain is derived from the SRE-ZBP protein (Attar, R. M. and Gilman, M. Z. (1992). Expression cloning of a novel zinc-finger protein that binds to the c-fos serum response element. Mol.Cell. Biol. 12, 2432-2443; Genbank accession number: M88579). SRE-ZBP is a member of the C_2H_2 class of zinc-finger proteins. It has seven tandem zinc fingers. Any one of these zinc-finger domains or any combination of two or more domains can be used to generate a DNA-binding domain with novel recognition specificity. Furthermore, because the general

structure and mode of DNA recognition is known for proteins of this class, DNA recognition can be directly modified if necessary.

Transcriptional activation domains

In one embodiment, the SH3 domain will be linked to a DNA binding domain and therefore tethered to DNA associated with the reporter gene. The SH3 domain will have a previously identified peptide ligand binding partner. The SH3 ligand (SSL) will be expressed in cells as part of a hybrid protein which also contains a transcriptional activation domain. Binding of the SSL-bearing hybrid protein to the SH3-bearing transcriptional activator fusion protein is detected by monitoring reporter gene expression. The SH3 -SSL interaction results in the transcriptional activation domain being brought to promoter DNA, stimulating reporter gene transcription. Such SH3 interaction will result in the increased expression of a gene whose production and quantitation can be easily determined.

In general, preferred transcriptional activation domains are compact in size, potent in activity, and non-toxic to cells. Additional features become relevant for particular applications. For example, for constructs to be used in one particular type of cell or tissue, activation domains with especially high activity in the target cell are used. In applications that involve the stable integration of the target gene into host cell DNA, activation domains that resist chromatin suppression are used.

SH3 domains

SH3 domains for use in the fusion proteins comprise all or part of an SH3-containing protein, including at least the SH3 domain of interest or ligand binding region thereof. Typically, the SH3 domain of our fusion proteins will be at least about 40 amino acids in length, and may contain up to all or substantially all of the amino acid sequence of the SH3-bearing protein or a portion sufficient to satisfy the functional criteria of an SH3 domain as described herein. The extents of the domain may be defined by homology to other SH3 domains, reference to previously characterized SH3 domains or function. In one embodiment, the inclusion of domains other than those generally regarded as necessary for SH3 function may be included if the additional domain(s) is hypothesized to allosterically or otherwise influence the behavior of the SH3 domain. In other embodiments, inclusion of protein domains that are not generally regarded as influencing, either allosterically or otherwise, the behaviour of the SH3 domain may be included to stabilize the SH3 fold, enhance expression or provide a means of identifying or manipulating the protein (*e.g.* fusion to glutathione-S-transferase or other domains that aid in identification/purification, epitope tags, *etc*).

SH3 ligand domains

SH3 ligand domains for use in the fusion proteins will generally be as described above. In a cell free implementation, SH3 ligands may be synthetic or non-synthetic peptides, peptoids or small molecules identified from libraries, cell broths, natural extracts, *etc.*

optional domains

Compounds introduced into cells may be labile and therefore not present for long periods of time once introduced into the cell. It would be advantageous to have all components present in the cell at the time the cells are exposed to potential inhibitors of SH3 binding partners. We therefore have made constructs that consist of a steroid hormone ligand binding domain (HBD) fused to the SH3 ligand/activation domain protein. The regulatory properties of the steroid hormone ligand binding domain (Picard et al, Cell (1988) 54, 1073-1080; Eilers et al, Nature (1989) 340, 66-68) allows the post-translational interaction of the SH3 binding partner pair. The steroid hormone ligand binding domain serves as an autonomous regulatory domain: upon fusion to a heterologous protein, the hormone binding domain subjects the chimeric protein to hormonal control. This regulatory property is due to the hormone-reversable interactions of the HBD with heat shock proteins. Hence, the DNA binding domain/SH3 domain and the SH3 ligand/HBD/transcriptional activation domain can be expressed in cells containing the reporter gene. Only with the addition of steroid hormone will the SH3 binding partners interact. In the absence of hormone, the SH3 ligand/HBD/transcriptional activation domain fusion protein will not be competent to interact with the DNA bound SH3 binding partner.

The HBD is an important feature of the system described. Inhibitors of SH3-dependent protein-protein interactions could exert their effect in two ways; compete with the ligand/HBD/activation domain for the binding partner or displace the ligand/HBD/activation domain), which has complexed with the binding partner. Mechanistically, it will be easier to block than displace binding partners. Potential inhibitors can be introduced into cells prior to complex formation (hormone addition) to block complex formation.

In addition to the SH3 ligand and activation domain and HBD, other domains that might be part of this hybrid protein include an antibody recognition (epitope) sequence (for detection using antibodies to confirm that the fusion protein is made and that the level of expression is not affected by compounds tested for SH3 binding partner inhibition) and a nuclear localization sequence.

4 **Assembly of constructs**

DNA sequences encoding the DNA-binding or transcriptional activating domain(s) is joined to DNA encoding the SH3 domain(s) or SH3 ligand domain(s), as the case may be. These sequences are joined such that they constitute a single open reading frame that can be translated in cells into a single polypeptide harboring all intended domains. The order and arrangement of the domains within the polypeptide can vary.

5 **Introduction into cells & selection of transfectants.**

These constructs may be introduced into cells by any convenient means. It may be most convenient to use simple transfection procedures. The constructs are combined onto a single plasmid vector together with a selectable marker gene, or they may reside on any number of individual plasmids which could be introduced by co-transfection. All three constructs may be combined into a single viral vector, or single or multiple plasmids may be administered by direct DNA injection or in conjunction with lipid carriers.

Engineered cells and cell lines may be prepared by introducing into them the reporter gene of interest linked to the abovementioned expression elements and to a selectable marker, usually a constitutively expressed selectable marker. Introduction of DNA at this and other steps of the preparation of the cells may be carried out using any convenient procedures, including for example calcium phosphate transfection or lipofection. Read through transcription, *e.g.* from the promoter of the constitutively expressed selectable marker, may result in a detectable or even high basal level of expression for the reporter gene. [Therefore, cells will be selected that exhibit low or undetectable levels of reporter gene expression. An expression vector that produces a DNA binding domain/activation domain fusion protein will be introduced into the cell lines to select for cells with high levels of induced reporter gene expression.

6 **Setting up & running the assays**

The SH3-inhibitor indicator cell line will be cultured in standard tissue culture media containing the drugs necessary to select for cells which stabily retain the genes described above. Cells will be cultured in standard tissue culture dishes, which includes multidishes and microwell plates. Cells will be exposed to hormone and the transcription induced due to SH3 binding partner interactions assayed using standard colorometric or flourescence assays.

Test compounds will be added to the culture media to assay the effect on SH3 binding. Cells at this time may be bathed in tissue culture media (with or without

serum) or a balanced salt solution. The compounds to be tested may be cell permeable and therefore added directly to cells or it may be necessary to first permeable the cells using streptolysin O, tetanolysin or another cell permeabilizing agent. Cells will be exposed to hormone (to induce SH3 binding partner interactions). After transcription is induced, cells or media(if a reporter gene such as SEAP is used,see above) will be harvested and an extract prepared to assay reporter gene expression.

7 how to interpret the results

The SH3 fusion protein and SSL fusion protein are capable of forming a complex with each other. The cells express the reporter gene unless a substance is present which binds to the relevant SH3 domain or otherwise blocks the SH3-mediated interaction required for transcription of the reporter gene. In this assay, the cells are cultured or maintained in a suitable culture medium to establish a base-line for expression of the reporter gene. The test substance is added to the culture medium and the ability of the test substance to inhibit expression of the reporter gene is measured. If the level of reporter gene expression is reduced in the presence of the test substance, the test substance is a putative SH3 blocker with respect to the SH3 domain involved in transcriptional control. There may be compounds that inhibit reporter gene expression by non-specific mechanisms, for example by having a general inhibitory effect on cellular transcription or translation. Because of this possibility, compounds will also be assayed for their effect on reporter gene expression when reporter gene transcription is not dependent on SH3 binding pair interactions. For this purpose, an indicator cell line will be utilized wherein the transcriptional activation domain is brought directly to DNA by the DNA binding domain (by expressing in cells a DNA binding domain/transcriptional activation domain fusion). The fusion protein will also contain a hormone binding domain (HBD) so that the only difference between the two indicator cell lines is the dependence upon SH3 binding pair interaction for reporter gene expression.

8 Pharmaceutical applications

By virtue of its capacity to inhibit protein-protein interactions required for cellular events of pharmacologic importance, SH3 binders and blockers identified by the method ("SH3 inhibitor") of this invention may be used in pharmaceutical compositions and methods for treatment or prevention in a mammal in need thereof.

Mammals include rodents such as mice, rats and guinea pigs as well as dogs, cats, horses, cattle, sheep, non-human primates and humans.

The preferred method of such treatment or prevention is by administering to a mammal an effective amount of the SH3 inhibitor to prevent, alleviate or cure said disease or disorder. Such effective amounts can be readily determined by evaluating the compounds of this invention in conventional assays well-known in the art, including assays described herein.

Therapeutic/Prophylactic Administration & Pharmaceutical Compositions

The invention provides methods of treating, preventing and/or alleviating the symptoms and/or severity of a disease or disorder referred to above by administration to a subject of an SH3 inhibitor in an amount effective therefor. The subject will be an animal, including but not limited to animals such as cows, pigs, chickens, *etc.*, and is preferably a mammal, and most preferably human.

Various delivery systems are known and can be used to administer the SH3 inhibitor, *e.g.*, encapsulation in liposomes, microparticles, microcapsules, *etc.* One mode of delivery of interest is via pulmonary administration, as detailed more fully *infra*. Other methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural and oral routes. The SH3 inhibitor may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.*, oral mucosa, rectal and intestinal mucosa, *etc.*) and may be administered together with other biologically active agents. Administration can be systemic or local. For treatment or prophylaxis of nasal, bronchial or pulmonary conditions, preferred routes of administration are oral, nasal or via a bronchial aerosol or nebulizer.

In specific embodiments, it may thus be desirable to administer the SH3 inhibitor locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, by injection, by means of a catheter, by means of a suppository, or by means of a skin patch or implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers.

This invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically (or prophylactically) effective amount of the SH3 inhibitor, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The carrier and composition can be sterile. The formulation should suit the mode of administration.

The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. The composition can be a liquid

solution, suspension, emulsion, tablet, pill, capsule, sustained release formulation, or powder. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, *etc.*

In a specific embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic to ease pain at the side of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

Administration to an individual of an effective amount of the SH3 inhibitor can also be accomplished topically by administering the compound(s) directly to the affected area of the skin of the individual. For this purpose, the SH3 inhibitor is administered or applied in a composition including a pharmacologically acceptable topical carrier, such as a gel, an ointment, a lotion, or a cream, which includes, without limitation, such carriers as water, glycerol, alcohol, propylene glycol, fatty alcohols, triglycerides, fatty acid esters, or mineral oils.

Other topical carriers include liquid petroleum, isopropyl palmitate, polyethylene glycol, ethanol (95%), polyoxyethylene monolaurate (5%) in water, or sodium lauryl sulfate (5%) in water. Other materials such as anti-oxidants, humectants, viscosity stabilizers, and similar agents may be added as necessary.

In addition, in certain instances, it is expected that the SH3 inhibitor may be disposed within devices placed upon, in, or under the skin. Such devices include patches, implants, and injections which release the compound into the skin, by either passive or active release mechanisms.

Materials and methods for producing the various formulations are well known in the art [see *e.g.* US Patent Nos. 5,182,293 and 4,837,311 (tablets, capsules and other oral formulations as well as intravenous formulations)].

The effective dose of the SH3 inhibitor will typically be in the range of about 0.01 to about 50 mg/kgs, preferably about 0.1 to about 10 mg/kg of mammalian

body weight, administered in single or multiple doses. Generally, the SH3 inhibitor may be administered to patients in need of such treatment in a daily dose range of about 1 to about 2000 mg per patient.

The amount of the SH3 inhibitor which will be effective in the treatment or prevention of a particular disorder or condition will depend on the nature of the disorder or condition, and can be determined by standard clinical techniques. In addition, *in vitro* or *in vivo* assays may optionally be employed to help identify optimal dosage ranges. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems. The precise dosage level of the SH3 inhibitor, as the active component(s), should be determined by the attending physician or other health care provider and will depend upon well known factors, including route of administration, and the age, body weight, sex and general health of the individual; the nature, severity and clinical stage of the disease; and the use (or not) of concomitant therapies.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceutical or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

Pulmonary Administration

In one embodiment of this invention, the SH3 inhibitor is administered by pulmonary administration, *e.g.* via aerosolization. This route of administration may be particularly useful for treatment or prophylaxis of bronchial or pulmonary infection or tumors.

Pulmonary administration can be accomplished, for example, using any of various delivery devices known in the art (see *e.g.*, Newman, S.P., 1984, in Aerosols and the Lung , Clarke and Davia (eds.), Butterworths, London, England, pp. 197-224; PCT Publication No. WO 92/16192 dated October 1, 1992; PCT Publication No. WO 91/08760 dated June 27, 1991; NTIS Patent Application 7-504-047 filed April 3, 1990 by Roosdorp and Crystal), including but not limited to nebulizers, metered dose inhalers, and powder inhalers. Various delivery devices are commercially available and can be employed, *e.g.*, Ultravent nebulizer (Mallinckrodt, Inc., St. Louis, Missouri); Acorn II nebulizer (Marquest Medical Products, Englewood, Colorado), Ventolin metered dose inhaler (Glaxo Inc., Research Triangle Park, North Carolina); Spinhaler powder inhaler (Fisons Corp.,

Bedford, Massachusetts) or Turbohaler (Astra). Such devices typically entail the use of formulations suitable for dispensing from such a device, in which a propellant material may be present.

Ultrasonic nebulizers tend to be more efficient than jet nebulizers in producing an aerosol of respirable size from a liquid (Smith and Spino, "Pharmacokinetics of Drugs in Cystic Fibrosis," Consensus Conference, Clinical Outcomes for Evaluation of New CF Therapies, Rockville, Maryland, December 10-11, 1992, Cystic Fibrosis Foundation).

A nebulizer may be used to produce aerosol particles, or any of various physiologically acceptable inert gases may be used as an aerosolizing agent. Other components such as physiologically acceptable surfactants (*e.g.*, glycerides), excipients (*e.g.*, lactose), carriers, and diluents may also be included.

This invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are intended to fall within the the scope of the appended claims.

Various patents, patent applications and publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

Examples

EXAMPLE 1 Phage Display Library Construction

Reagents and Strains. T4 DNA ligase and T4 DNA polymerase were from Biorad. DNA sequencing was performed using a USB Sequenase kit as described by the manufacturer. Glutathione Sepharose 4B was purchased from Pharmacia. Oligonucleotides were synthesized with an Applied Biosystems xxx and purified using OPC columns. Peptides were synthesized using an Applied Biosystems synthesizer and purified by HPLC. Panning was performed using Costar 24 well (#3424) polystyrene dishes.

Phage display library construction. M13 bacteriophage, containing an amber triplet (TAG) at position 20 of gene III DNA, was propagated in SupE E. coli host JM101. Single stranded DNA was prepared as described (Zoller and Smith, Meth. Enzymol., 154:329 (1982)). Oligonucleotides were designed such that inserted sequences would be placed after position 20 of gene III. The oligonucleotides were also designed such that recombinants would lack the gene III amber codon. The E. coli

strain JS5, which does not suppress amber mutations, was used as the host to select for recombinants.

Phosphorylation of oligonucleotide, annealing, synthesis of complementary DNA and electroporation was as described (Sambrook et al., Mol. Col. (Laboratory Manual) CSHL (1989); Zoller and Smith *Supra*). Four phage display libraries were prepared. The oligonucleotides used for the synthesis of the libraries were:

First library (X_6) (SEQ ID NO: 160)
(5')taaacaactttcaacagtcccKNN CNN CNN CNN CNN CNN
cccagcggagtgagaata(3') (SEQ ID NO.:161)

Second library (X_6PPIP) (SEQ ID NO: 3)
(5')gttttgctaaacaactttcaacagtcccgggtattggtgg KNN KNN KNN KNN KNN KNN
ttcagcggagtgagaatagaaagg(3') (SEQ ID NO.:162)

Third library (RSLRPLX_6) (SEQ ID NO: 65)
(5')gttttgctaaacaactttcaacagt YNN YNN YNN YNN YNN YNN
caaggggcggagtgagcgttcagcggagtgagaatagaaagg(3') (SEQ ID NO.:163)

Fourth library (PPPYPPX_6) (SEQ ID NO: 100)
(5')gttttgctaaacaactttcaacagt YNN YNN YNN YNN YNN YNN
cggtggatatggtggtggttcagcggagtgagaatagaaagg(3') (SEQ ID NO.:164)

Underlined nucleotides represent complementary M13 sequences. Sequences inserted into gene III are not underlined. Nucleotides used to create random amino acid sequences are denoted N, K and Y (N=A,G,C+T; K=G+T; Y=C+A). The other gene III-inserted sequences are described in the text.

Amplification

Bacteriophage were amplified in 500 mls of 2XYT media containing 10 mls of stationary phase JS5 cells at 37°C, 250 rpm for 6 hrs. The supernatant was cleared twice by centrifugation to remove bacteria (4000 rpm for 15 min in a Sorvall H4000 rotor at 4°C) and phage precipitated by adjusting the supernatant to 0.5M NaCl/4% PEG.

EXAMPLE 2 Isolation of Phage with Affinity for Selected SH3 Domains

A. Affinity Selection. Phage containing peptides with affinity for SH3 domains were isolated by a method known as panning (Parmley and Smith, Gene,

73:305-318 (1988)). Wells (24 dishes) were coated first using 10 µg of GST or GST-SH3 protein in 100 mM $NaCO_3$, pH 8.5 o/n at 4°C and then for 30 min at room temperature using a 20 mg/ml BSA (in 100 mM $NaCO_3$, pH 8.5) blocking buffer. Phage (1.5×10^{11} infectious particles) were panned in 250 µl of 50 mM Tris HCl, pH 7.5/150 mM NaCl containing 0.5% BSA and 0.05% Tween 20 o/n at 4°C. Wells were washed 3 X using TBS/0.5% BSA/0.05% Tween 20 at room temperature. Bound phage were eluted using 200 µl of 0.2M Glycine, pH 2.1/0.5% BSA/0.05% Tween 20 and neutralized using 50 µl 1M Tris base. Eluted phage used again for panning were first amplified in 20 mls 2XYT media containing 0.4 mls of stationary phase JS5 cells and purified/concentrated as described above.

B. Isolation and Sequencing of Phage Selected by the Src SH3 Domain.

Degenerate oligonucleotides were designed so that a phage library could be prepared containing six residues of random amino acid sequences inserted into the bacteriophage M13 gene III polypeptide. These random amino acid sequences were inserted into gene III, bracketed by glycine residues, so that the polypeptide might be flexible and presented more efficiently as part of a gene III fusion. A library of 2.5×10^9 recombinants were generated.

GST and GST-Src SH3 were each immobilized on polystyrene and binding selection was performed. After six cycles of selection/amplification, GST Src SH3-dependent phage enrichment was observed; the phage titer was 100-fold higher with GST-Src SH3, compared to a binding assay using GST. After eight cycles (approximately 1000 fold enrichment), ten bacteriophage were isolated and sequenced. Each phage had the same nucleotide sequence, encoding a proline-rich peptide insert. In addition, a point mutation in flanking sequence (GGG-->AGG) had occurred, converting a glycine residue to arginine:

--GXXXXXXG-- (SEQ ID NO.: 157) starting library

--RSLPPIPG-- (SEQ ID NO.: 158) GST-Src SH3 selected

The phage displayed Src SH3 selected peptide is interesting as it has some similarity to part of a protein, termed 3BP1, identified by Cicchetti et al., who screened a lgt11 cDNA expression library using the SH3 domain of Abl. This group also found that Src's SH3 domain could bind to 3BP1. The region of 3BP1 sequence important for SH3 binding is proline rich, as is the peptide identified by phage display.

Curiously, the isolation of phage containing the proline-rich sequence required six cycles of binding selection. Phage enrichment is usually observed after 2-3 cycles. That six cycles were required, and only phage with the same nucleotide sequence identified, suggests that the phage selected by GST-Src SH3 were not present in the original library used for screening but instead generated during binding selection. The

single point mutation, which changed the glycine residue to arginine, probably occurred during the amplification of phage and was important for Src SH3 recognition. These results support the view that the Src SH3 recognition sequence is proline-rich, at least seven amino acid residues in length and perhaps preferring an arginine residue as part of the recognition sequence.

To expand our understanding of Src SH3 domain binding preferences, a new phage display library was prepared ($2X10^9$ recombinants) that contained six random amino acid residues inserted into gene III, adjacent to the proline-rich sequences identified in the first screen: --XXXXXXPPIP-- (SEQ ID NO: 3), referred to as the X_6PPIP sequence.

The X_6PPIP library (SEQ ID NO: 3) was screened, using the Src SH3 domain (as a GST fusion) and with GST. GST-Src SH3 retained 10x more phage than GST after one cycle of selection suggesting the enrichment of certain phage. After three cycles of binding selection (approximately a 1000 fold enrichment) phage were isolated and sequenced. The amino acid sequences (SEQ ID NOS: 4 to 64) identified by screening the library using the Src SH3 domain are shown in Figure 1.

With respect to the six randomized residues upstream of PPIP, at positions 1, 2 and 3, no particular amino acid preferences were observed. In contrast, at position 4; arginine was preferred, at position 5; proline and at position 6; leucine. That arginine is preferred at the fourth position is consistent with the notion that the point mutation observed with the isolate from the first library (SEQ ID NO: 2) (a glycine to arginine change) was important for Src SH3 selection. We also note that the inclusion of the amino acids PPIP as part of the gene III displayed residues probably is important for Src SH3 recognition as phage enrichment was observed after only one round of panning.

To further explore SH3 binding preferences, the X_6 PPIP phage display library (SEQ ID NO: 3) was also screened using the Fyn, Lyn, PI3K, Abl domains (as GST fusions) or using Grb-2 protein (which contains two SH3 domains). Four distinct ligand classes were identified (see Figure 1):

Src, Fyn SH3 domains -	XXXRPL(PPIP) (SEQ ID NO: 165)
Lyn, PI3K SH3 domains -	RXXRPL(PPIP) (SEQ ID NO: 166)
Abl SH3 domain -	PPPYPP(PPIP) (SEQ ID NO: 167)
Grb-2 SH3 domains -	RLYRPL(PPIP) (SEQ ID NO: 168)

To more precisely define the ligand preferences of the SH3 domains listed above, three new libraries were prepared; RSLRPLXXXXXX, to further define Src, Fyn, Lyn and PI3K SH3 binding preferences (SEQ ID NO.: 65); PPPYPPXXXXXX, to further define Abl SH3 binding preferences (SEQ ID NO.: 100) and RLYRPLXXXXXX, to further define Grb-2 SH3 binding preferences (SEQ ID NO.: 109). Bacteriophage

were selected from these new libraries using GST-SH3 fusions or Grb-2, and after four rounds of panning, phage were isolated and sequenced. Again, specific binding preferences were observed (Figure 2).

Src, Lyn SH3 domains -(RSLRPL)PPLPXP (SEQ ID NO: 169)
PI3K SH3 domain - (RSLRPL)PPLPPP (SEQ ID NO: 170)
Grb-2 SH3 domains - (RLYRPL)PPLPXX (SEQ ID NO: 171)
Abl SH3 domain - (PPPYPP)PPI/VPXX (SEQ ID NOS: 172 and 173)
Fyn SH3 domain - (RSLRPL)PPI/VPXX (SEQ ID NOS: 174 and 175)

Randomizing six amino acid residues at a time, we have identified specific peptide ligands recognized by the Src, Fyn, Lyn, PI3K, Abl and Grb-2 SH3 domains. With the above studies, we have randomized a "window of SH3 binding recognition" that totals twelve residues in length. To determine whether additional (flanking) sequences contribute to binding, two new libraries were prepared that were designed to be useful for studying Src, Fyn, Lyn and PI3K SH3 binding preferences. These new phage displayed peptides include "core sequences", essentially a consensus sequence derived from the results of the screens described above. In addition, five randomized residues were placed either upstream (--XXXXXRPLPPLPPP--, SEQ ID NO.: 133) or downstream (-RSLRPLPPLPXXXXX–, SEQ ID NO.: 114) of the core sequences. These two libraries were screened to identify additional SH3 binding peptides using the Src, Lyn and PI3K SH3 domains. Specific binding preferences were observed. The selected sequences are shown in Figure 4A and Figure 4B.

To determine whether these new sequences are important determinants for specific binding, two phage-displayed peptides, isolated from the X_5RPLPPLPPP (SEQ ID NO: 133) library using the Lyn and Src SH3 domains (Lyn selected: --PWWLDRPLPPLPPP--, (SEQ ID NO.: 141) and VPLARRPLPPLPPP--, (SEQ ID NO.: 149)) were subjected to BIAcore® analysis (see above). GST-Lyn and GST-Src SH3 domains were immobilized on a Biosensor chip and the interaction between SH3 domain and phage-displayed peptides were measured (Figure 5). Binding selectivity was clearly observed. Phage selected using the Lyn SH3 domain bound more efficiently to the Lyn SH3 domain and the phage selected using the Src SH3 domain bound more efficiently to the Src SH3 domain.

C. Isolation and Sequencing of Phage Selected by the Abl SH3 Domain.

They X_6PPIP (SEQ ID NO: 3) phage display library was also used to more precisely define Abl SH3 binding preferences. As shown in Figure 3, the sequences selected by the Abl SH3 domain are distinct from those identified by the Src SH3 domain. Because the Abl selected sequences are distinct, a new library, PPPYPPX_6 (SEQ ID NO: 100) was prepared and screened to identify more precisely the ligands

preferred by the Abl SH3 domain. The results of the $PPPYPPX_6$ library screen are shown in Figure 2 (SEQ ID NO: 101 to 108).

D. Summary of Results

We have constructed phage display libraries to generate a rich source of structurally diverse peptides to identify high affinity ligands for the Src, Fyn, Lyn, PI3K, Grb-2 and Abl SH3 domains. We observe that although peptides selected by these SH3 domains are proline-rich, they are distinct. In this study, inroads towards understanding SH3 binding specificity was first obtained with the fortuitous isolation of a Src SH3 selected phage displayed peptide and the finding that the Src SH3 recognition sequence was proline-rich and at minimum seven residues in length. This information was important for the design of a phage display library XXXXXXPPIP, (SEQ ID NO.: 3) that when screened, yielded insight as to Src, Fyn, Lyn, PI3K, Grb-2 and Abl SH3 binding preferences (Figure 1). We continued this strategy by preparing and screening additional phage display libraries that contained either five or six random residues adjacent to sequences that helped to constitute SH3 recognition sequences. This approach has led to the identification of distinct SH3 binding preferences. Based on information obtained from screening seven phage display libraries, we have defined what we term the SH3 ligand core (Figure 1 and 2) and SH3 ligand flanking sequences (Figures 4A and 4B).

The Src, Fyn, Lyn, PI3K, Abl and Grb-2 SH3 ligand core sequences were defined using the libraries XXXXXXPPIP (SEQ ID NO.: 3), RSLRPLXXXXXX (SEQ ID NO.: 65), PPPYPPXXXXXX (SEQ ID NO.: 100) and RLYRPLXXXXXX (SEQ ID NO.: 109). That specific flanking sequences could be identified was demonstrated using the Src, Lyn and PI3K SH3 domains and the libraries XXXXXRPLPPLPPP (SEQ ID NO.: 133) and RSLRPLPPLPXXXXX (SEQ ID NO.: 114). The claim of SH3 specific ligands is based on the unique sequences identified for a particular SH3 binding domain (Figures 1, 2, 4A and 4B) as well as the specificity observed with respect to binding affinity (Figures 3 and 5).

The core peptides identified in this report define the minimal sequence requirements for efficient recognition by the Src, Fyn, Lyn, PI3K, Grb-2 and Abl SH3 domains. When using Src, Lyn or PI3K SH3 domains to screen phage display libraries that containing random residues either upstream or downstream of core residues, we observe additional amino acid preferences (see figures).

The single letter amino acid codes used herein are standard and represent amino acids including the following:

R = Arginine;	S = Serine;	L = Leucine;	P = Proline;
Y = Tyrosine;	G = Glycine;	I = Isoleucine;	V = Valine;
A = Alanine;	W = Tryptophan;	X = any amino acid.	

Example 3 Assessment of Binding of Phage-Selected Peptides to *Src* and *Abl* SH3 Domains

The *Src* and *Abl* SH3 domains (*Src* amino acids 81-140 and *Abl* amino acids 84-138) were expressed as carboxy-terminal fusions to glutathione-S-transferase (amino acids 1-226 (Smith, D.B. and Johnson, K.S., Gene 67:31-40 (1988)) in *Escherichia coli* and purified to >95% homogeneity by glutathione affinity chromatography. GST-*Src* and GST-*Abl* fusion proteins were covalently immobilized via solvent-exposed lysines to a BIAcore Biosensor Chip CM-5 (Pharmacia Biosensor, Piscataway, NJ) using standard N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide hydrochloride/N-hydroxysuccinimide chemistry (O'Shannessy, D.J. et al., Analyt. Biochem. 205:132-136 (1992)). The resulting GST-*Src* and GST-*Abl* biosensor chips were docked into a BIAcore Biosensor (Pharmacia Biosensor, Piscataway, NJ) and the biospecific interaction between immobilized SH3 domain and phage-displayed peptides measured by surface plasmon resonance (Malmqvist, M., Nature 361:186-187 (1993); Jonsson, U. and Malmqvist, M., Advances in Biosensors, JAI Press Ltd., London, 1992, pp. 291-336; Jonsson, U. et al., BioTechniques 11(5):620-627 (1991)). Interaction between receptor (SH3) and ligand (phage-displayed peptide) is indicated by an increase in RU value (positive inflection in the plane of the ordinate). Assays were run at 25°C in phosphate buffered saline, pH 7.4, containing 0.05% Surfactant P20 ((Pharmacia Biosensor, Piscataway, NJ) at a flow rate of 5 μL min^{-1}. 10^{13} pfu/ml of *Src*-selected phage displaying the peptide sequence RSSRPLPPIP (SEQ ID NO.: 35 or *Abl*-selected phage displaying the peptide sequence PPPYPPPPIP (SEQ ID NO.: 49) were used to measure binding.

Results of these assays are represented graphically in Figures 3A-3D. As is evident, *Src*-selected phage bound specifically to immobilized GST-*Src* SH3 (Figure 3A) and did not bind to immobilized GST-*Abl* SH3 (Figure 3B). As also shown, *Abl*-selected phage bound specifically to immobilized GST-*Abl* SH3 (Figure 3D) and did not bind to immobilized GST-*Src* SH3 (Figure 3C). Thus, these results demonstrate the binding specificity and selectivity of a ligand to selected SH3 domain. Specifically, the results show that a peptide ligand selected on the basis of its ability to bind *Src* binds *Src* SH3 but not *Abl* SH3 and, conversely, that a peptide ligand selected on the basis of its ability to bind *Abl* binds *Abl* SH3 but not *Src* SH3.

EXAMPLE 4 Association of Src Protein With Tyrosine Phosphorylated Proteins *in vivo*

Several proteins have been reported to coprecipitate with v-Src from extracts of v-Src-transformed cells (Kanner, S.B. et al., EMBO J. 10:1689-1698 (1991); Koch, C.A., Mol. Cell. Biol. 12:1366-1374 (1992). To further analyze the involvement of the SH3 domain in the association of these proteins with Src, we examined the ability of the Src SH3-binding peptides to inhibit the coprecipitation of Src with tyrosine phosphorylated proteins. We also examined the ability of a high affinity Src SH2-binding peptide (EPQpYEEIPI (SEQ ID NO: 176), denoted here as YEEI) to block SH2 domain-dependent interactions. In this assay, Balb/c 3T3 and SRD-3T3 cells (V-Src transformed 3T3 cells) were lysed in RIPA buffer and the Src protein was immunoprecipitated with a monoclonal antibody to the amino-terminus of Src in the presence or absence of the SH2 and SH3 binding peptides. For comparison, the SH3 binding proteins were also isolated by affinity isolation using GST-Src-SH3. The samples were transferred to nitrocellulose, and the blot was probed with a monoclonal antibody to phosphotyrosine. Two major tyrosine phosphorylated proteins co-migrated with v-Src from SRD 3T3 cells, one of which co-migrated with the 62 kDa tyrosine-phosphorylated protein that bound to the Src SH3 domain, and the other protein of Mr 130,000, which is most likely the 130,000 kDa protein previously detected in v-Src immunoprecipitates. (Koch, C.A., Mol. Cell. Biol. 12:1366-1374 (1992)).

The binding of Src to the 62 kDa protein was reduced significantly by incubation of the extracts with the "Src-pro" peptide or the phosphorylated YEEI peptide, but not the poly-proline nor the unphosphorylated YEEI peptides, suggesting that both the SH2 and SH3 domains of Src are required for stable association of this protein with Src. We have been unable to determine whether this protein is the p62 protein identified in the total cell SH3 binding proteins since the p62 antibody is not sensitive enough to recognize the small amounts of the 62 kDa protein that associate with Src under these conditions. The "Src-pro" peptide and the pYEEI peptide could only block 50% of p130 binding to Src, whereas 90% inhibition of p130 binding was observed when both peptides were used, suggesting that either the SH2 or the SH3 domain is sufficient for some degree of p130 association with Src.

EXAMPLE 5

SSLs may be identified based on their ability to inhibit formation of an SH3/ligand complex. A known ligand, preferably a peptide, is immobilized to the surface of the sensor material or chip as defined for the instrument(*e.g.* see*BIAcore® System Manual*, BIAcore Biosensor AB, Uppsalla, Sweden, 1994) using known

chemistries (*BIAapplications Handbook*, BIAcore Biosensor AB, Uppsalla, Sweden, 1994). The ability of a potential SSL to bind to a known SH3 domain is inferred from its ability to inhibit the formation of SH3 domain/ligand complex as measured in the presence and absence of various concentrations of SSLs using surface plasmon resonance technologies.

For example, 360 resonance units (RU) of a Grb-2 SH3-ligand (Ac-EVPVPPPVPPRRRPGGK-NH_2) were covalently immobilized to a Biosensor Chip CM5 (Pharmacia Biosensor) through the e-amino group of the lysyl residue. Immobilizations were performed at 25°C with a flow rate of 5 ul min^{-1} in HBS running buffer (10 mM HEPES pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.05% Surfactant P20 (Pharmacia Biosensor)) as follows: 30 uL injection of 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride / 0.05 M N-hydroxy-succinimide to create reactive linkers; 30 uL injection 0.1 mM peptide in HBS to permit covalent linking; 30 uL injection of 1.0 M ethanolamine hydrochloride for 5 minutes to block unreacted N-hydroxy-succinimide esters. To remove non-covalently bound peptide and stabilize the hydrogel, the chip was cycled through ten consecutive regeneration loops consisting of 2 minutes exposure to 6 M guanidine HCl (pH 7.0); 5 minutes exposure to HBS (flow rate of 5 uL min^{-1}). Identical wash cycles were used to regenerate the chip between sensorgram.

Assays were initiated by combining 1uM Grb-2(P204L) SH3 domain with various concentrations (100, 50, 25, 10, 5, 2.5, 1 or 0 uM) of SOS-derived peptide (Ac-EVPVPPPVPPRRRPGGK-NH2) or test peptide (Ac-SHRLYRPLPPLPGGK-NH2) and pre-incubated at 25°C for >30 minutes. Assays were performed in HBS at 25°C with a flow rate of 3 ul min^{-1} in PBS/imidazole (1.2 mM KH2PO4, 8.1 mM Na2HPO4.7H20, 138 mM NaCl, 0.5 mM MgCl2.6H2O, 2.7 mM KCl, 0.9 mM CaCl2, 80 mM immidazole, pH 7.4). Complex formed between the SH3 domain and the immobilized ligand was measured using the BIAcore Biosensor® (Pharmacia Biosensor, Piscataway, NJ) for each concentration of competitor and the IC50 calculated. See Figure ZX.

Alternatively, SSLs may be identified based on their ability to bind to SH3 domains, again, as measured using BIAcore® and allied technologies. For example, an SSL (or SH3 domain) is immobilized to the surface of the sensor material or chip as defined for the instrument(*e.g.* see*BIAcore® System Manual*, BIAcore Biosensor AB, Uppsalla, Sweden, 1994) using known chemistries (*BIAapplications Handbook*, BIAcore Biosensor AB, Uppsalla, Sweden, 1994) as above. The ability of the immobilized SSL (or SH3 domain) to bind to an SH3 domain (or SSL) is directly measured using surface plasmon resonance technologies. See Rickles et al, EMBO *supra*.

Claims

1. A method for identifying a peptide comprising an amino acid sequence not present in a naturally occurring protein, which peptide is a ligand for a protein of interest containing at least one SH3 domain, the method comprising:

(a) contacting the SH3 domain with a mixture of one or more candidate peptides under conditions permitting a ligand to bind to an SH3 domain to form an SH3 domain-ligand complex,

(b) separating unbound peptides from any SH3 domain-ligand complexes so formed,

(c) dissociating the complexed peptide ligand(s) from the SH3 domain-ligand complexes,

(d) enriching for the selected peptide(s) by re-contacting the dissociated peptides with the SH3 domain under conditions permitting a ligand to bind to an SH3 domain to form a complex, and,

(e) detecting which if any candidate ligands bound to the SH3 domain.

2. A method of claim 1 which further comprises

(a) dissociating any candidate peptide ligands from SH3-ligand complexes formed during the enrichment step, and

(b) identifying the amino acid sequences of the dissociated peptide ligands.

3. A method for identifying a peptide comprising an amino acid sequence not present in a naturally occurring protein, which peptide is a ligand for a protein of interest containing at least one SH3 domain, other than Abl or phosphatidylinositol 3-kinase, which method comprises:

(a) combining an SH3 domain of interest with one or more candidate peptide ligands under conditions permitting a ligand to bind to an SH3 domain to form an SH3-ligand complex,

(b) detecting which if any candidate ligands bound to the SH3 domain.

4. A method for identifying a peptide ligand for a protein of interest containing at least one SH3 domain, other than Abl, Src, Grb2, Nck or phosphatidylinositol 3-kinase which method comprises:

(a) combining an SH3 domain of interest with one or more candidate peptide ligands under conditions permitting a ligand to bind to an SH3 domain to form an SH3-ligand complex,

(b) detecting which if any candidate ligands bound to the SH3 domain.

5. A method of any of claims 1 through 4 wherein the SH3 domain is provided in the form of a fusion protein.

6. A method of any of claims 1 through 5 wherein the mixture of peptides is provided in the form of a library of synthetic peptides or in the form of a phage library displaying the various peptides.

7. A method of any of claims 1 through 6 which further comprises contacting the ligand so identified with an SH3 domain from a second protein under conditions permitting the formation of an SH3 domain-ligand complex and determining the presence or extent of binding of the ligand to the SH3 domain from the second protein.

8. A peptide ligand for an SH3 domain of interest identified by the method of any of claims 1 through 7.

9. A peptide ligand for an SH3 domain of interest comprising an amino acid sequence not present in a naturally occurring protein, which was identified by the method of any of claims 1 through 7.

10. A peptide ligand specific for an SH3 domain of interest, identified by the method of of any of claims 1 through 7, which binds to the SH3 domain against which it was selected with a Kd value of less than half of the Kd with which it binds to a different SH3 domain..

11. A peptide ligand of claim 8, 9 or 10 containing at least seven amino acid residues.

12. A peptide ligand of claim 11 containing at least 10 amino acid residues.

13. A peptide ligand of claim 12 containing at least 12 amino acid residues.

14. A peptide ligand of claim 13 containing at least 14 amino acid residues.

15. A peptide ligand of claim 14 containing at least 15 amino acid residues.

16. A peptide comprising an amino acid sequence not present in a naturally occurring protein, which peptide is a ligand capable of binding selectively to an SH3 domain from src, fyn, lyn, yes, txk, tsk, btk, NADPH oxidase p47-phox or p67-phox, or crk.

17. A peptide ligand capable of binding selectively to an SH3 domain from the group consisting of (1) members of the src-family protein tyrosine kinases (including Lyn, Lck, Hck, Fgr, and Yes), (2) Sprk, a threonine/serine protein kinase, (3) Tsk, (4) Btk, (5) Txk and other Tec family members, (6) Vav, (7) GTPase Activating Protein (GAP), (8) p40, p47, and p67 proteins of the neutrophil oxidase complex, (9) Crk, and (10) phospholipase C gamma.

18. A DNA sequence comprising a nucleotide sequence encoding a peptide ligand of any of claims 8 through 17.

19. A method for identifying a substance which inhibits the interaction of proteins mediated by an SH3 domain of interest, which method comprises:

(a) providing (i) an SH3 domain of interest, (ii) a specific SH3 ligand (SSL) comprising a peptide of claim 8, 9, 10, 11, 12, 13, 14, 15, 16, or 17 which is capable of selectively binding to the SH3 domain of interest to form an SH3-SSL complex and (iii) a composition comprising a test substance to be evaluated;

(b) incubating the SH3 domain and the SSL under conditions permitting the formation of an SH3-SSL complex, in the presence and absence of the test substance; and

(c) measuring the ability of the test substance to inhibit the formation or persistence of the SH3-SSL complex.

20. A method for identifying a substance which inhibits the interaction of proteins mediated by an SH3 domain of interest, which method comprises:

(a) providing (i) an SH3 domain of interest, (ii) a specific SH3 ligand (SSL) which is capable of selectively binding to the SH3 domain of interest to form an SH3-SSL complex but comprises an amino acid sequence not present in a

naturally occurring protein and (iii) a composition comprising a test substance to be evaluated;

(b) incubating the SH3 domain and the SSL under conditions permitting the formation of an SH3-SSL complex, in the presence and absence of the test substance; and

(c) measuring the ability of the test substance to inhibit the formation or persistence of the SH3-SSL complex.

21. A method for identifying a substance which inhibits the interaction of proteins mediated by an SH3 domain of interest, which method comprises:

(a) providing (i) an SH3 domain of interest other than an SH3 domain from Grb-2 or Fyn, (ii) a specific SH3 ligand (SSL) which is capable of selectively binding to the SH3 domain of interest to form an SH3-SSL complex and (iii) a composition comprising a test substance to be evaluated;

(b) incubating the SH3 domain and the SSL under conditions permitting the formation of an SH3-SSL complex, in the presence and absence of the test substance; and

(c) measuring the ability of the test substance to inhibit the formation or persistence of the SH3-SSL complex.

22. A method of claim 19, 20 or 21 in which the SH3 domain is provided in the form of a fusion protein containing the SH3 domain linked to heterologous peptide sequence.

23. A method of claim 19, 20 or 21 in which the SSL is a peptide of at least 7 amino acid residues.

24. An inhibitor of protein-protein interactions mediated by an SH3 domain of interest identified by the method of any of claims 19 through 23.

25. A genetically engineered cell containing and capable of expressing:
(a) a reporter gene,
(b) a first DNA sequence encoding a DNA-binding fusion protein containing a DNA-binding domain and a first SH3 binding partner, and

(c) a second DNA sequence encoding a transcriptional activating fusion protein containing a transcriptional activating domain and a second SH3 binding partner capable of binding to the first SH3 binding partner,

wherein

(i) the reporter gene is under the expression control of a DNA element to which the DNA-binding fusion protein is capable of binding,

(ii) the transcriptional activating fusion protein is capable of binding to the DNA-binding fusion protein, and

(iii) expression of the reporter gene is measurably activated under conditions permitting the binding of the DNA-binding fusion protein to the DNA element and to the transcriptional activating fusion protein.

26. A genetically engineered cell of claim 25, wherein:

(a) the first SH3 binding partner is a peptide sequence comprising the amino acid sequence of an SH3 domain of interest, and the second SH3 binding partner is a peptide sequence comprising a specific ligand for the SH3 domain of interest; or

(b) the first SH3 binding partner is a peptide sequence comprising a specific ligand for the SH3 domain of interest, and the second SH3 binding partner is a peptide sequence comprising the amino acid sequence of an SH3 domain of interest.

27. A method for identifying an inhibitor of interactions between proteins mediated by one or more SH3 domains which comprises:

(a) culturing cells of claim 25 or 26 in an appropriate culture medium permitting growth of the cells and under suitable conditions permitting detectable expression of the reporter gene,

(b) adding a test material to the culture medium, and

(c) identifying any decrease in the expression of the reporter gene.

28. A method of claim 27 wherein the test material comprises a sample of: a microbial broth; cellular extract; conditioned medium from cell lines or from host cells transformed with genetic libraries; a collection of one or more synthetic compounds; or a combinatorial library of compounds.

29. An inhibitor identified by the method of claim 27 or 28.

--- *ala* GLY xxx xxx xxx xxx xxx xxx GLY *thr val* ---

arg ser leu pro pro ile pro 1st library isolate (src SH3)

xxx xxx xxx xxx xxx xxx pro pro ile pro	X_6PPIP library
ser leu ala arg pro leu	
thr ser met arg pro leu	
lys ser glu arg pro leu	
thr leu gly arg pro leu	
ser ile ala arg pro leu	
ala pro arg ile pro leu	
leu his arg arg ala leu	
tyr asn his arg ser leu	
leu arg gln arg pro leu	SRC
pro ala ser arg pro leu	
ala arg asp arg pro leu	
pro arg ser arg pro leu	
phe val ser arg pro leu	
asn lys gly arg ser leu	
asp arg leu arg pro leu	
leu ala asn arg glu leu	
pro thr arg arg pro leu	
ile gln his arg leu leu	
arg leu leu lys pro leu	
gln ser arg arg ser leu	
ala lys arg ala pro leu	
arg leu leu lys pro leu	FYN
thr gly gly arg pro leu	
his his ile arg pro leu	
ser tyr pro arg pro leu	
his ser thr arg ala leu	
pro thr his arg ile leu	
arg asn phe leu ile leu	
arg val ser tyr asn leu	
pro val leu his ser leu	
arg tyr leu arg pro leu	
arg ser ser arg pro leu	LYN
val asn lys arg pro leu	
arg ser asn arg pro leu	
arg ala lys arg pro leu	
ser his arg his pro leu	
ala ser thr arg pro leu	
ala ser thr arg pro leu	
arg leu phe arg pro leu	
arg pro gln arg pro leu	PI3K
arg phe lys arg pro leu	
arg ala lys arg pro leu	
arg pro tyr arg pro leu	
arg ile pro arg pro leu	
arg ser leu arg pro leu	

Figure 1A/14

pro pro pro tyr pro pro		
pro pro pro tyr pro pro (4x)		
ala pro pro tyr pro pro (2x)		
ala pro his tyr pro pro		
pro pro pro tyr his pro		
pro pro ser tyr pro pro		ABL
pro pro ala tyr pro pro		
pro pro ala tyr pro pro		
ala pro asn tyr pro pro		
ala pro ser tyr ser pro		
ala pro ser tyr pro pro (2x)		
pro pro his tyr pro pro		
arg leu tyr arg pro leu (3x)		
arg leu tyr arg pro leu		
arg leu his arg pro leu		GRB2
arg his tyr his pro leu (3x)		
arg his his arg pro leu (4x)		
leu ser lys arg pro leu		
val his arg arg pro leu		
leu ser arg arg pro leu		
phe arg asp arg pro leu		
arg ala asn arg pro leu		
ile his thr arg pro leu		YES
leu his arg arg pro leu		
ile his thr arg pro leu		
leu ala arg arg pro leu		
lys ala lys arg pro leu		
leu ala arg arg pro leu		
ala glu pro tyr leu lys thr pro		
phe his phe arg pro pro		TSK
gly tyr asn lys pro pro		
pro tyr gly lys val pro		
arg glu ala tyr ser lys thr pro (3x)		
ala tyr gly lys pro pro		

Figure 1B/14

arg ser leu arg pro leu	xxx xxx xxx xxx xxx xxx
	pro pro leu pro gly ser
	pro pro leu pro arg pro
	pro pro leu pro pro pro
SRC	pro pro leu pro gly ser
	pro pro leu pro pro pro
	pro pro leu pro pro pro
	pro pro leu pro ser pro
	pro pro val pro pro pro
	pro pro leu pro asn gly
	pro pro leu pro phe asp
	pro pro pro pro phe pro
	pro pro leu pro val pro
	pro pro leu pro trp --- thr*
FYN	pro pro leu pro pro leu
	pro pro ile pro gly arg
	pro pro ile pro pro pro
	pro pro leu pro ala pro
	pro pro ile pro leu gly
	pro pro ile pro asp ser
	pro pro leu pro ala pro
	pro pro leu pro ser pro
	pro pro leu pro leu pro
	pro pro leu pro leu arg
LYN	pro pro leu pro leu arg
	pro ser leu pro phe pro
	pro pro leu pro pro pro
	pro pro leu pro leu leu
	pro pro leu pro gln leu
	pro pro leu pro ser ala
	met val ser leu val pro
	leu pro phe gly pro pro
	pro pro thr pro val-glu ser**
PI3K	pro pro leu pro pro pro
	pro pro leu pro pro pro
	pro pro leu pro pro pro
	pro pro leu pro pro pro
	pro pro leu pro pro pro
pro pro pro tyr pro pro	xxx xxx xxx xxx xxx xxx
	pro pro ile pro val arg
	pro pro val pro tyr ala
	pro pro ile pro ile leu
	pro pro val pro tyr pro
ABL	pro pro val pro arg ser
	pro pro val pro gly ala
	pro pro ile pro arg phe
	pro asp ile pro leu leu
arg leu tyr arg pro leu	xxx xxx xxx xxx xxx xxx
	pro pro leu pro trp pro (2x)
	pro pro leu pro ser arg
Grb-2	pro pro leu pro pro val
	pro pro leu pro ile thr

Figure 2/14

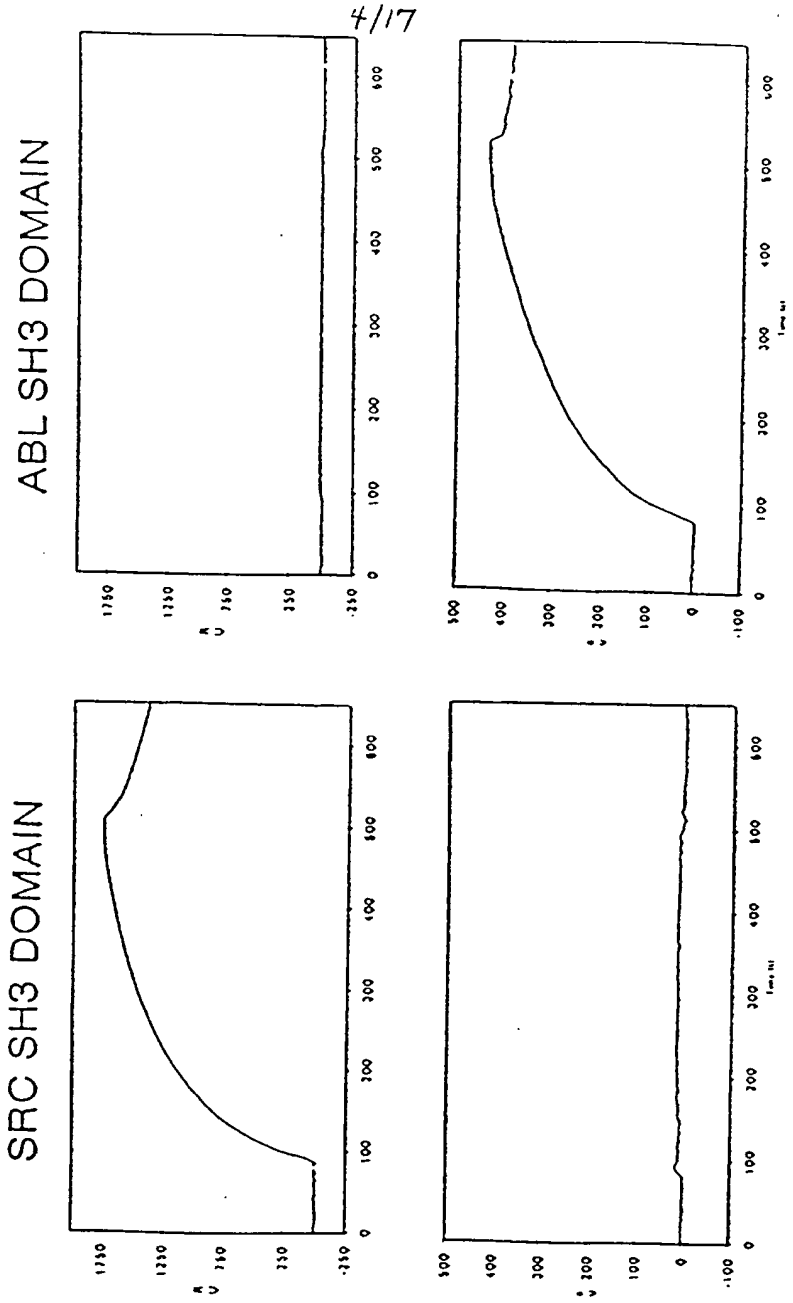

ABL SH3 DOMAIN
SRC SH3 DOMAIN
Phage-Displayed Sequence
RSSRPLPPIP
PPPYPPPPIP

Figure 3/14

xxx xxx xxx xxx xxx arg pro leu pro pro leu pro pro pro		
trp ser leu phe his		
ser trp leu ser arg	(2x)	
cys leu leu cys ser		
ile his leu arg thr		
pro his pro ala arg		
val pro leu ala arg	(2x)	(GST-SRC SH3)
val ser leu ala gln		
ile ser leu ala arg		
ser trp leu ala asn		
leu ser leu gln met		
ser ala pro ser arg		
leu arg trp ser ala		
lys ala leu ala ser		
thr thr leu ser ala		
ser tyr leu ser arg		(GST-FYN SH3)
leu ser pro met ser		
leu pro met ala arg		
ala ser phe ala ser		
val gly leu ser asn		
ile pro ile gly his		
ile ala leu ser thr		
phe phe leu ala ser		
leu asp arg ser arg		
arg trp ile thr ser		
gly leu gly ala arg		
asp asp leu ser gly		
asn pro leu leu arg		
ser trp leu ile ala		
arg ala leu pro pro		
trp ser ile arg ser		
tyr pro phe his gly		
pro trp leu leu arg		
ser trp leu ser val		(GST-LYN SH3)
ser trp leu gly ser		
pro trp trp leu asp		
leu thr met ala gly		
asp phe met ala ser		
asp trp tyr his ser		
leu ser phe ser asn		
met asp leu leu arg		
ala ala met val gly		
arg leu leu pro pro		
ser trp val ala cys		
pro trp pro val gly		
pro asn leu arg thr		
leu pro leu ser arg		
ser trp trp ser arg		
thr phe leu phe arg		
thr his pro ala arg		
phe pro val ala gln		

Figure 4A/14

pro ser pro arg ser (GST-YES SH3)
arg ser leu ser met
val leu leu ala leu
ser val leu ala arg
asn arg ala ala lys
ser ala leu ala arg
ser ser leu ala gln
thr pro leu gly arg
pro asn pro ser tyr
ser ala ser ala arg
ser trp met ala arg
ser ala leu ala arg
ser pro leu ala arg

xxx xxx xxx xxx xxx arg pro leu pro pro leu pro pro pro
cys leu asn cys leu (4x)
phe lys arg pro pro
leu leu ser gly pro (3x)
leu leu ser gly pro
phe lys arg pro phe
leu val leu pro arg (GST-PI3K SH3)
phe his arg pro asn
arg arg val pro pro
trp ser arg thr lys
arg val arg pro ser
arg tyr leu pro trp
cys cys leu thr thr
arg pro phe pro arg
arg val pro ile arg

Figure 4B/14

arg ser leu arg pro leu pro pro leu pro xxx xxx xxx xxx xxx

GST-LYN (9x)

leu pro pro arg leu
leu pro pro arg thr
leu pro pro arg his
leu pro pro pro his

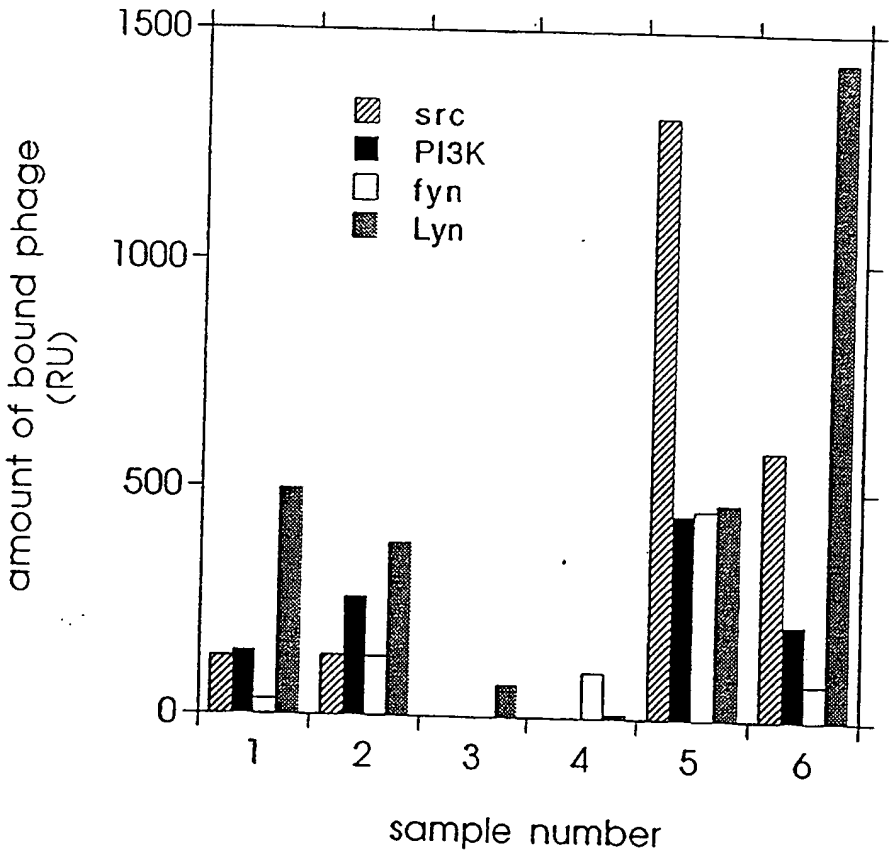
GST-PI3K

leu pro pro arg leu
phe pro arg tyr leu
pro ser ser ala pro
pro arg ala pro phe
pro arg pro pro phe
ser arg pro gly pro
pro arg pro ala tyr

GST-SRC

arg ser leu arg pro leu pro --- leu pro pro val pro ser leu
arg ser leu arg pro leu pro pro --- pro pro val pro ser leu
arg ser leu --- pro leu pro pro leu pro ala arg pro his pro
arg ser phe arg pro leu pro pro leu pro thr leu leu pro ser
arg ser leu arg pro leu pro pro leu pro thr pro pro leu his

Figure 5/14



(1) ser ile ala arg pro leu pro pro ile pro
(2) lys ser glu arg pro leu pro pro ile pro
(3) ser his arg his pro leu pro pro ile pro
(4) ala lys arg ala pro leu pro pro ile pro

(5) val pro leu ala arg arg pro leu pro pro leu pro pro pro
(6) pro trp trp leu asp arg pro leu pro pro leu pro pro pro

Figure 6/14

CLASS II +5

gly ala ala pro pro leu pro pro arg xxx xxx xxx xxx xxx

	asn pro pro arg leu
	asn lys val arg gly
	asn gln thr phe arg
	asn pro pro arg thr
(GST-SRC SH3)	asn arg leu his arg
	asn val pro arg leu
	asn arg leu his ala
	asn arg val lys leu
	asn arg val arg leu
	asn arg val val leu
	asn lys pro arg leu
	asn ile pro arg val
	asn pro pro arg lys
	asn lys pro arg leu
	asn pro pro arg thr
	asn lys ala arg leu
	asn his leu ser thr
	asn pro pro arg val
	asn pro arg pro ile
	asn arg val arg leu
(GST-FYN SH3)	asn ala val arg leu
	asn arg pro his phe
	asn pro lys his arg
	asn arg val arg leu
	asn pro pro arg phe
	asn ser arg pro leu
	asn pro ala arg leu
	asn thr pro arg leu
	asn pro arg pro leu
	asn pro pro arg thr
	asn pro thr pro val
	thr tyr gly thr thr
(GST-LYN SH3)	pro pro trp met met
	thr arg tyr ser leu
	pro pro trp met thr
	thr tyr leu leu pro
	thr tyr trp met pro
	pro thr phe met ala
	pro thr trp met thr
	thr tyr leu met pro
	pro ala trp met ala
	ser tyr leu leu pro
	thr tyr leu leu pro
	thr thr pro tyr thr
	pro ser trp met ser
	pro val thr trp ile
	pro ser phe met ser
	pro ala phe phe ser
	pro ser phe leu gln

Figure 7A/14

gly ala ala pro pro leu pro pro arg	xxx xxx xxx xxx xxx	
	asn arg pro arg val	
	asn pro pro arg ile	(2x)
	asn pro lys his arg	
	asn pro pro arg ser	
	asn pro lys his trp	
(GST-YES SH3)	asn pro val arg val	
	asn lys ala arg leu	
	asn lys his pro pro	
	asn pro pro arg lys	
	asn lys arg his leu	
	lys his leu ser ser	
	asn pro pro arg thr	(2x)
	asn pro pro arg val	
	asn lys leu arg val	
	pro pro leu arg ser	
	pro pro ile arg lys	
	pro pro arg pro leu	(2x)
(GST-PI3K SH3)	pro pro leu pro phe	
	pro pro lys pro ala	
	pro pro arg pro pro	
	pro pro arg pro ala	(2x)
	pro pro leu arg gln	
	arg ala leu arg leu	
	arg leu arg val gly	
	pro pro arg lys pro	
	pro pro arg thr pro	
	pro pro arg pro ser	
	pro pro arg pro phe	

Figure 7B/14

xxx xxx xxx xxx xxx xxx xxx xxx xxx xxx	10mer library
pro pro leu leu pro pro lys leu his ala (3x)	
ala pro pro val pro arg lys leu his pro (2x)	NH_2 Crk SH3
pro pro leu leu pro pro lys met his ser (4x)	
trp pro arg pro leu pro pro leu pro ala (1x)	
trp his arg pro leu pro pro leu pro pro (8x)	Lyn SH3
arg ser trp pro leu pro pro leu pro val (1x)	
arg pro leu pro leu pro pro val pro arg (7x)	
arg pro leu pro leu pro pro leu pro phe (2x)	Src SH3
arg pro leu pro ala pro pro pro arg phe (1x)	
pro trp arg pro leu pro ser ile pro leu (9x)	Fyn SH3

Figure 8/14

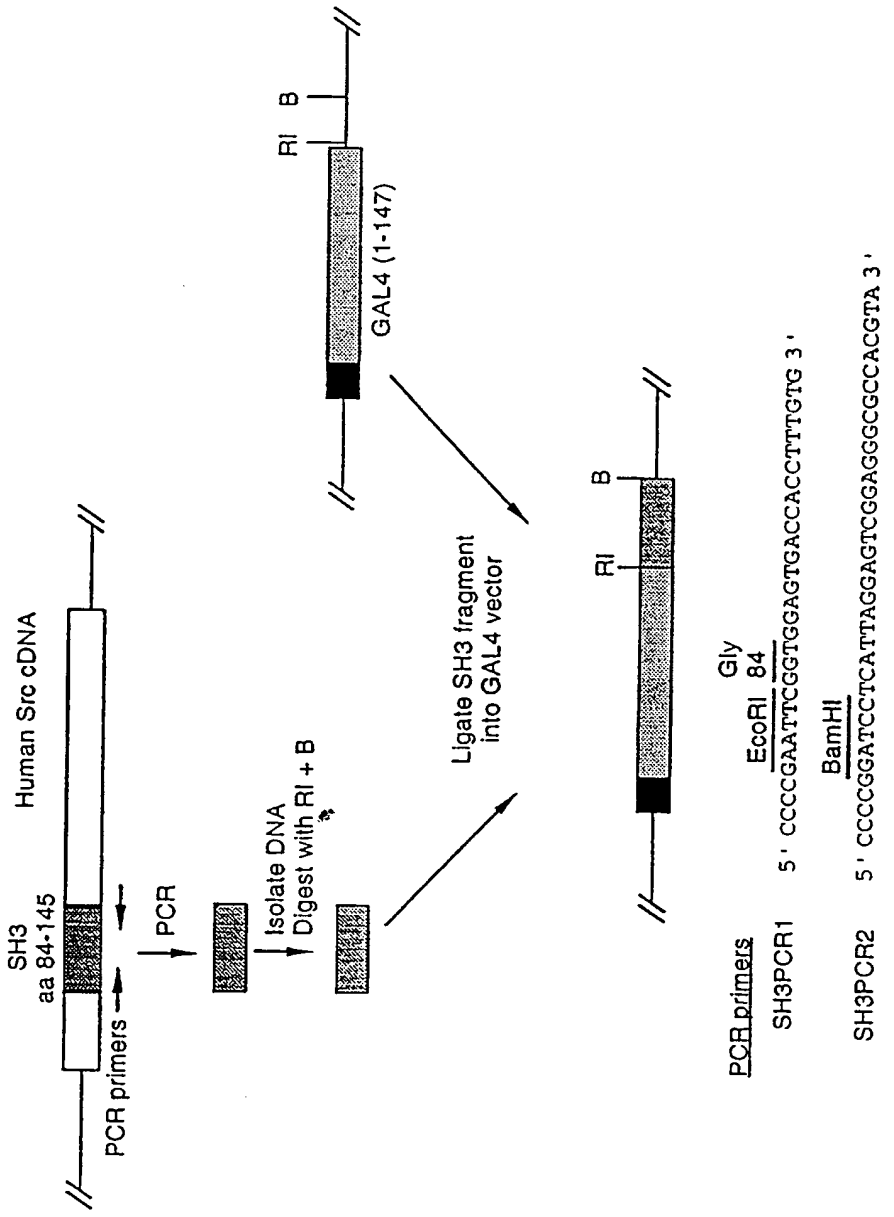

Human Src cDNA
SH3
aa 84-145
PCR primers
PCR
Isolate DNA
Digest with RI + B
Ligate SH3 fragment
into GAL4 vector
RI B
GAL4 (1-147)
RI B
Gly
EcoRI 84
PCR primers
SH3PCR1 5' CCCCGAATTCGGTGGAGTGACCACCTTTGTG 3'
BamHI
SH3PCR2 5' CCCCGGATCCTCATTAGGAGTCGGAGGGCGCCACGTA 3'

Figure 9/14

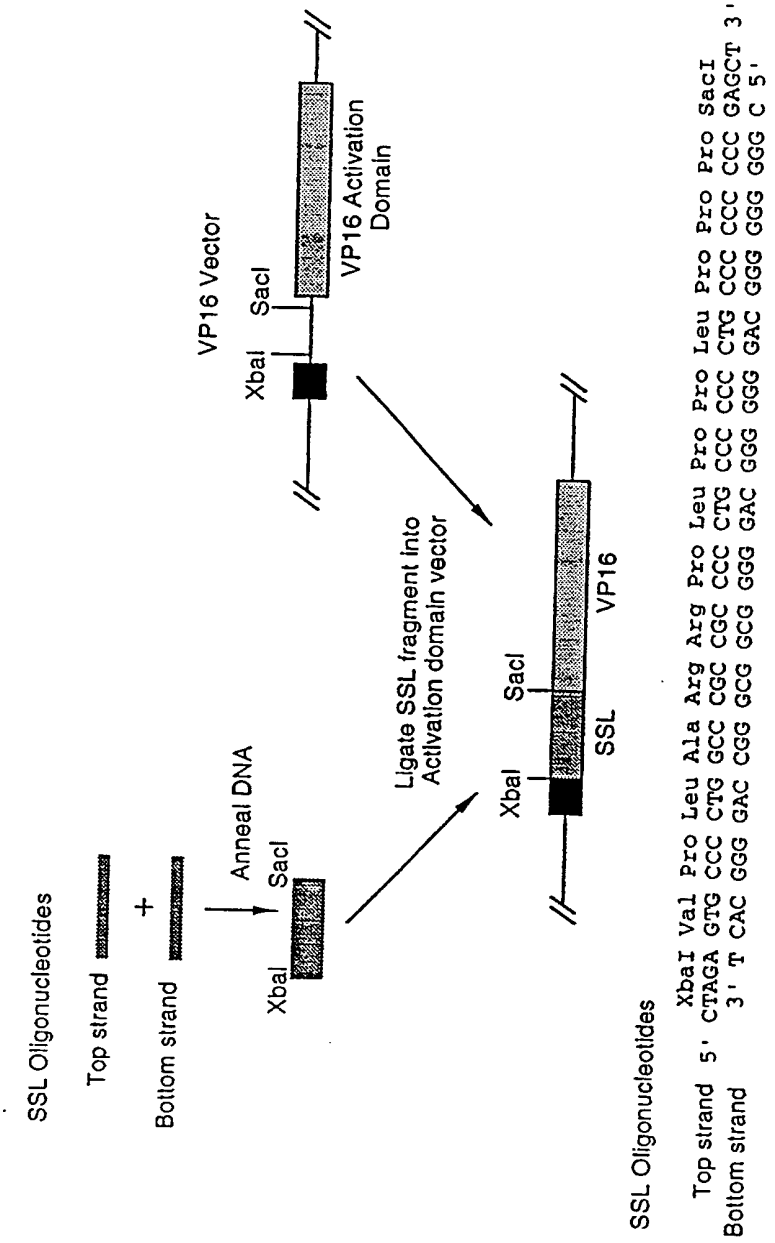

SSL Oligonucleotides
Top strand
+
Bottom strand
Anneal DNA
XbaI
SacI
Ligate SSL fragment into
Activation domain vector
VP16 Vector
XbaI
SacI
VP16 Activation
Domain
XbaI
SacI
SSL
VP16
SSL Oligonucleotides
XbaI Val Pro Leu Ala Arg Arg Pro Leu Pro Pro Leu Pro Pro Pro Pro SacI
Top strand 5' CTAGA GTG CCC CTG GCC CGC CGC CCC CTG CCC CCC CTG CCC CCC CCC GAGCT 3'
Bottom strand 3' T CAC GGG GAC CGG GCG GCG GGG GAC GGG GGG GAC GGG GGG GGG C 5'

Figure 10/14

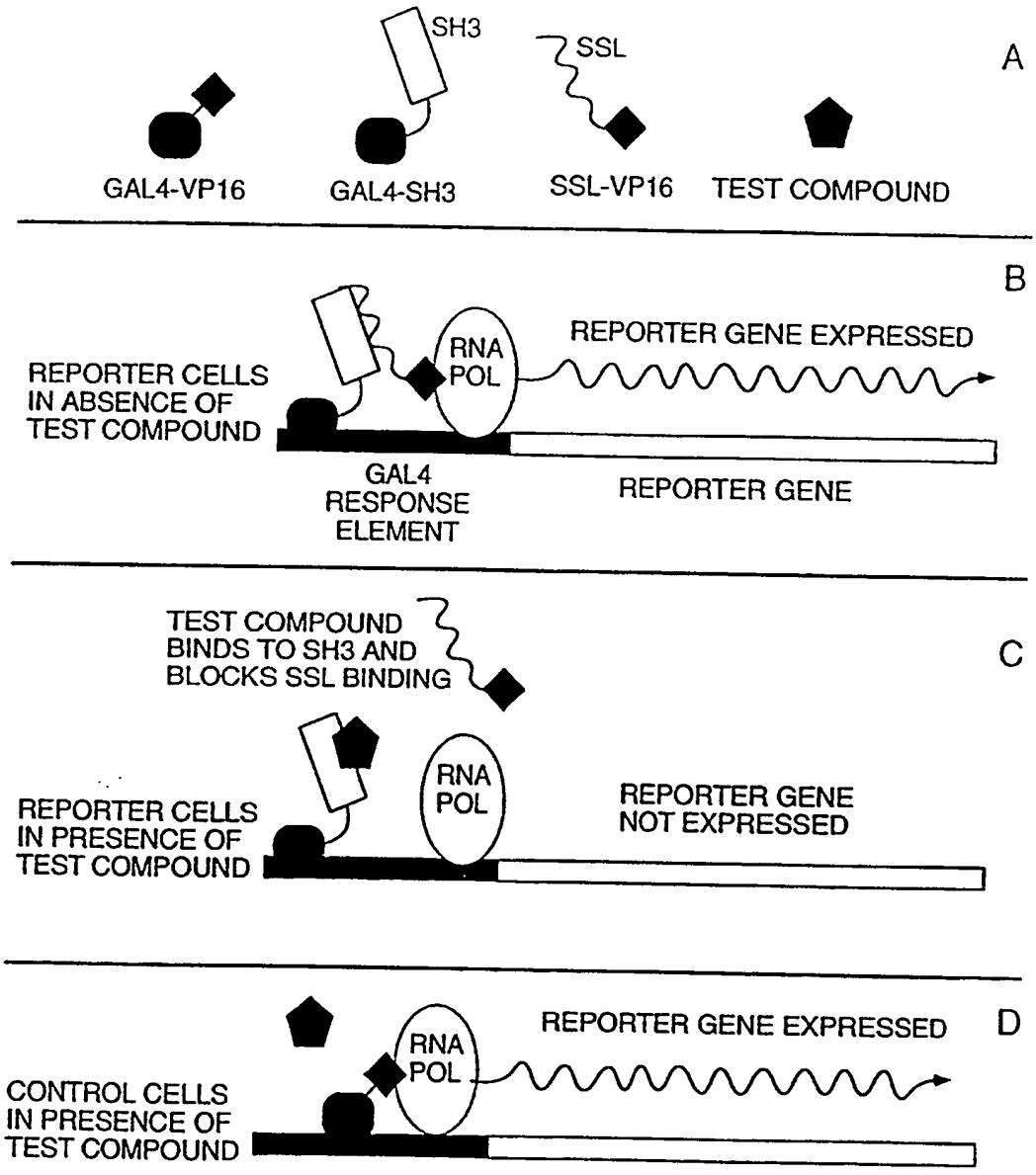

SH3
SSL
A
GAL4-VP16
GAL4-SH3
SSL-VP16
TEST COMPOUND
B
REPORTER CELLS IN ABSENCE OF TEST COMPOUND
RNA POL
REPORTER GENE EXPRESSED
GAL4 RESPONSE ELEMENT
REPORTER GENE
TEST COMPOUND BINDS TO SH3 AND BLOCKS SSL BINDING
C
REPORTER CELLS IN PRESENCE OF TEST COMPOUND
RNA POL
REPORTER GENE NOT EXPRESSED
CONTROL CELLS IN PRESENCE OF TEST COMPOUND
RNA POL
REPORTER GENE EXPRESSED
D

Figure 11/14

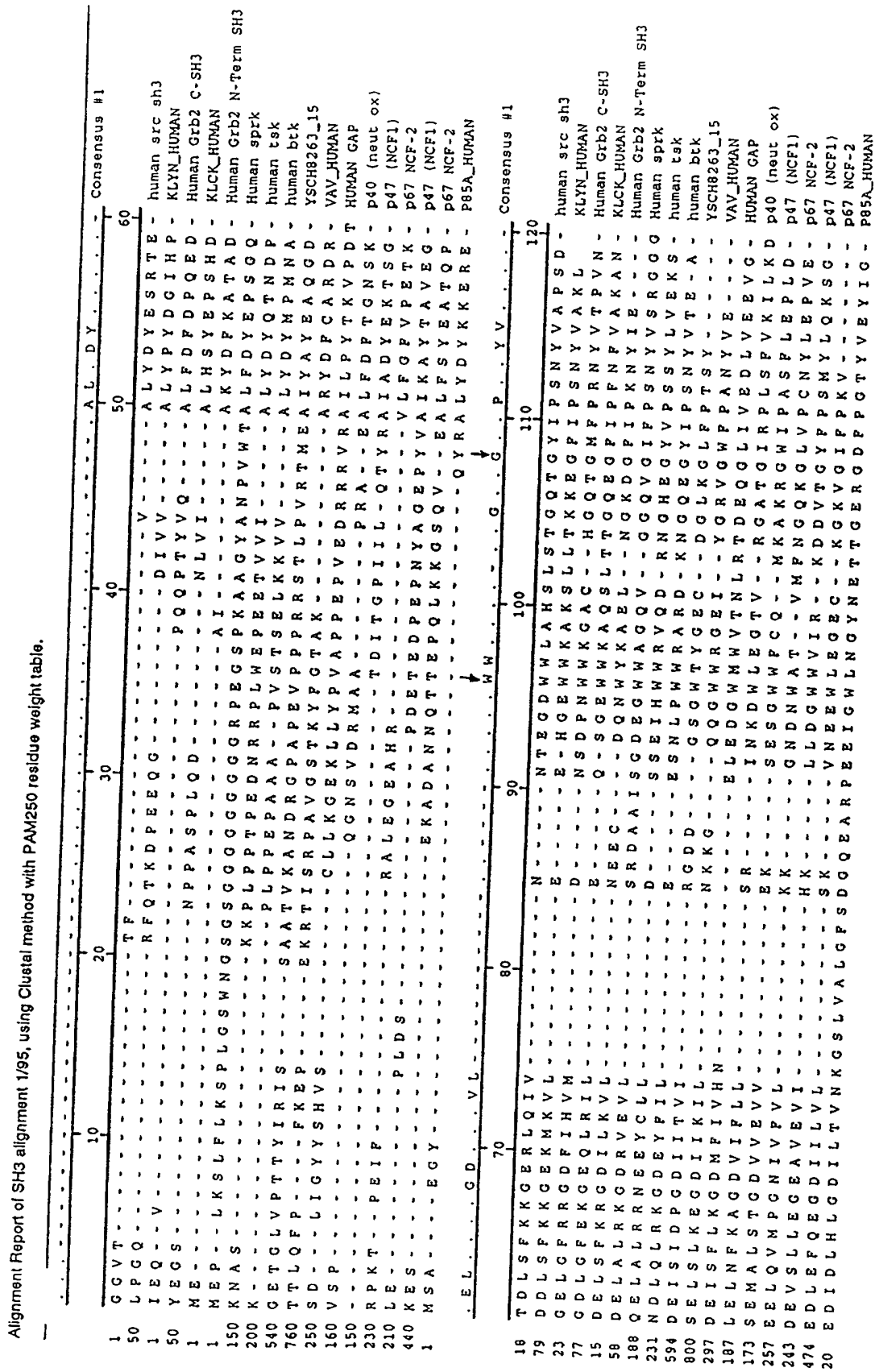

Alignment Report of SH3 alignment 1/95, using Clustal method with PAM250 residue weight table.
Consensus #1
human src sh3
KLYN_HUMAN
Human Grb2 C-SH3
KLCK_HUMAN
Human Grb2 N-Term SH3
Human sprk
human tsk
human btk
YSCH8263_15
VAV_HUMAN
HUMAN GAP
p40 (neut ox)
p47 (NCF1)
p67 NCF-2
P85A_HUMAN

Figure 12/14

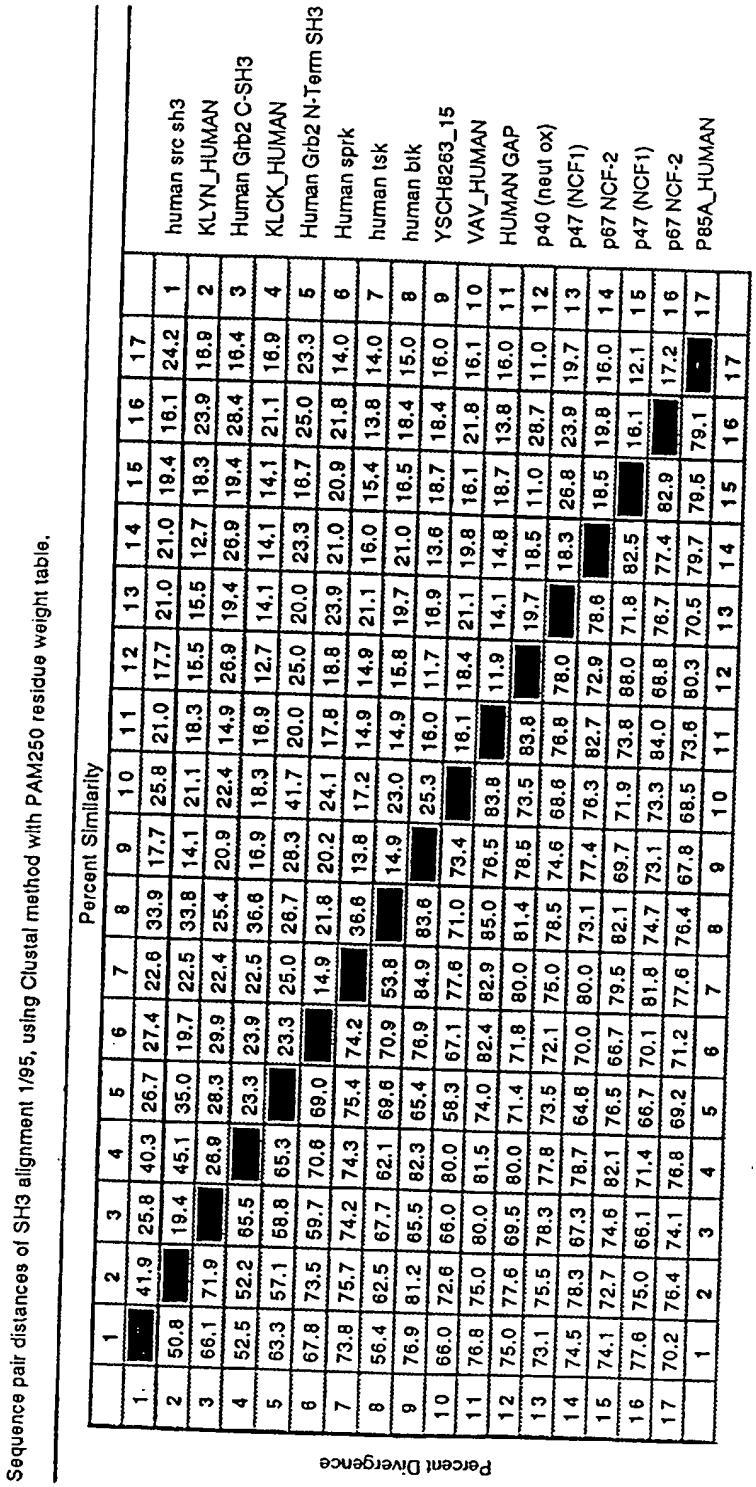

Sequence pair distances of SH3 alignment 1/95, using Clustal method with PAM250 residue weight table.
Percent Similarity
Percent Divergence
human src sh3
KLYN_HUMAN
Human Grb2 C-SH3
KLCK_HUMAN
Human Grb2 N-Term SH3
Human sprk
human tsk
human btk
YSCH8263_15
VAV_HUMAN
HUMAN GAP
p40 (neut ox)
p47 (NCF1)
p67 NCF-2
p47 (NCF1)
p67 NCF-2
P85A_HUMAN

Figure 13/14

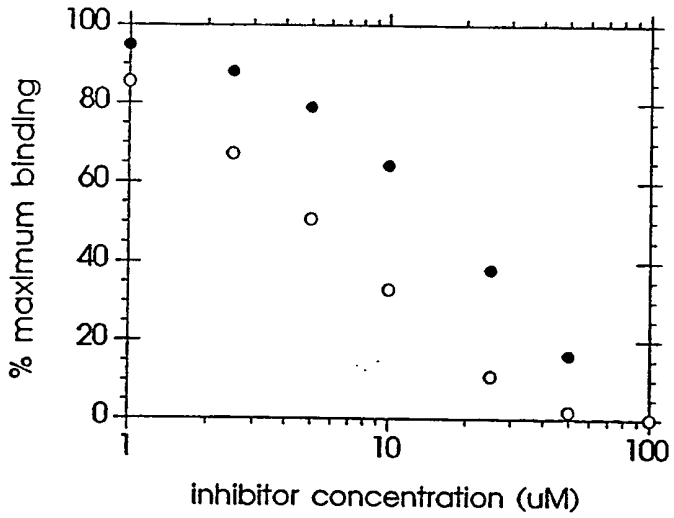

100
80
60
40
20
0
% maximum binding
1
10
100
inhibitor concentration (uM)

Figure 14/14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/03208

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07K 4/00; C12N 1/00, 5/00, 15/09, 15/12; C12Q 1/68; G01N 33/566

US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/5, 6, 7.1, 7.8, 29, 172.3, 240.1, 243; 436/501; 530/ 300, 326-329; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS and CAS ONLINE:
SH3(5A)(ligand# or peptide# or mimetic#)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A X Y	Journal of the American Chemical Society, Vol. 115, No. 26, issued 29 December 1993, JK CHEN ET AL, "Biased combinatorial libraries: Novel ligands for the SH3 domain of phosphatidylinositol 3-kinase", pages 12591-12592. See entire article.	1-2, 7, 19-23, 25-28 24, 29 3-6, 8-18
A, P	Journal of Organic Chemistry, Vol. 59, No.10, issued 20 May 1994, FA GOMEZ ET AL, "Affinity capillary electrophoresis: Insights into the binding of SH3 domains by peptides derived from an SH3-binding protein", pages 2885-2886. See entire article.	1-18

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

05 JUNE 1995

Date of mailing of the international search report

16 JUN 1995

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

DAVID SAUNDERS

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/03208

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	Cell, Vol. 76, No. 5, issued 11 March 1994, H YU ET AL, "Structural basis for the binding of proline-rich peptides to SH3 domains", pages 933-945. See page 934, col. 2-page 935, col. 1.	1-2, 19-23, 25-28
X, P		3-4, 6-11, 16-18, 24, 29
Y, P		5, 12-15
A, P	Journal of Biological Chemistry, Vol. 269, No. 39, issued 30 September 1994, AB SPARKS ET AL, "Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries", pages 23853-23856. See abstract and Figs. 1 and 2.	25-28
X, P		1-3, 5-18 24, 29
Y, P		4, 19-23
A, P	Journal of Biological Chemistry, Vol. 269, No. 39, issued 30 September 1994, C CHEADLE ET AL , "Identification of a Src SH3 domain binding motif by screening a random phage display library", pages 24034-24039. See abstract and Table I.	19-23, 25-28
X, P		1-3, 5-18, 24, 29
Y, P		4
A, P	EMBO Journal , Vol. 13, No. 23, issued 01 December 1994, RJ RICKLES ET AL, "Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries", pages 5598-5604. See abstract and Figs. 1 and 2.	19-23, 25-28
X, P		1-13, 16-18, 24, 29
Y, P		14-15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/03208

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/03208

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/5, 6, 7.1, 7.8, 29, 172.3, 240.1, 243; 436/501; 530/ 300, 326-329; 536/23.5

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

I. Claims 1-17 and 19-24, drawn to methods of identifying peptide ligands and peptide compounds thus identified.

II. Claim 18, drawn to a DNA sequence.

III. Claims 25-29, drawn to genetically engineered cells and their use in identifying inhibitors.
The DNA sequence of Group II has no common core structure with that of the peptides of Group I.

The genetically engineered cells of Group III contain no DNA sequence in common with that of Group II. These cells are not used in the methods of Group I. Accordingly, the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.